The languages of central and southern Philippines

Daniel Kaufman

To appear in: Antoinette Schapper and Sander Adelaar (eds.)
Oxford Handbook of Western Austronesian Languages. Oxford: Oxford University Press

1. Introduction

This chapter provides a typological overview of the languages of the Central and Southern Philippines (henceforth, CSP languages). Despite not forming a discrete phylogenetic group, the CSP languages share certain morphosyntactic retentions from Proto Malayo-Polynesian which make them a useful unit for typological generalizations. Like other Philippine languages, almost all the CSP languages maintain the full PMP voice system. On the other hand, the voice system mutates in interesting ways in the southernmost CSP languages, specifically, in the Bilic languages, covered here, and Sama languages, covered in Kaufman (this volume).

The languages within the scope of this chapter are those of the Greater Central Philippine subgroup (see Blust 1991 and Zorc, Lobel and Hall this volume), Kalamian (consisting of Agutaynen and Calamian Tagbanwa) and the Bilic subgroup (consisting of Tboli, Blaan and Tiruray). All these subgroups are argued to belong to a larger Philippine family by Blust (2019). Several members of the putative Philippine group are located outside of the Philippines, namely, the Sangiric, Minahasan and Gorontalic subgroups. These are excluded here on geographical grounds but are covered in Van den Berg and Mead (this volume). It should be noted that these languages have been influenced by distinct contact scenarios over the last several centuries, which have made them diverge morphosyntactically from their more northern relatives.[1] The Bilic languages, especially outside of Tboli, are still not sufficiently documented. Tboli thus serves here as a representative of this subgroup for present purposes with brief mention of Tiruray. Giangan, a language often classified as Bilic, is argued by Zorc (2019) to be an independent branch of a higher level subgroup. Giangan and Inati, potentially important witnesses, are too sparsely documented to be discussed here.

A series of work from the 1970s through the 1990s (Gallman 1983, Burton 1996, Elkins 1986, Savage 1986, Fleischman 1981) greatly improved our understanding of the languages of Mindanao and their interrelations. Two landmark dissertations cited frequently here, McFarland (1974) and Zorc (1977), provide comprehensive overviews of the Bikol and Bisayan languages, respectively, and include a wealth of comparative data on their morphological and syntactic structure. Gallman (1983), Burton (1996) and especially Pallesen (1985), show how contact effects have given shape to the vocabulary, phonological history and typology of several regions within the southern Philippines.

[1] I have reglossed the functional morphology in many of the examples here so that the terminology employed is as uniform as possible throughout. I do not mean to impose a particular analysis on the data by the use of "nominative" and "genitive" case, nor do I mean to imply that all forms glossed as "actor voice" are syntactically identical across languages. I transcribe examples of nasal substitution (triggered by the PMP prefixes *paŋ-/maŋ-) with deleted consonants in square brackets, e.g. *maŋ-[k]uːha*. I have also aimed to represent all the data presented here in a broad IPA transcription to avoid confusion across orthographies, although I maintain the symbol <y> for the palatal glide, as opposed to IPA [j]. Finally, any numbered examples whose language is not specified in the first line are Tagalog.

2. Phonology

2.1 Segment inventories

Vowel inventories in the CSP zone are relatively simple. The Central Philippine languages typically either preserve the Proto-Austronesian four vowel system (*i, *u, *a, *ə) or conflate it to a three-vowel system by merging *ə with one or more of the other vowels. In the languages of Mindanao, *ə is often preserved as a high central vowel (ɨ), and this was clearly the case in the not so distant past for many of the Central Philippine subgroups, as well. In rare cases, the inherited vowel inventory has been expanded in complex ways (e.g. Tboli, with its seven-vowel system, Porter 1977, Forsberg 1992).

Several languages have developed an allophonic relationship between the high vowels and their mid counterparts. In Tagalog, a generally word-final process of vowel lowering turns *i* and *u* into *e* and *o*, respectively. Kapampangan of the Central Luzon group (outside the purview of this chapter) has innovated a new set of mid vowels not from lowering of high vowels but rather through monophthongization of **aj* > *e* and **aw* > *o*, but this is vanishingly rare in the CSP zone. A large monophthongization zone begins just southeast of the CSP languages in Sulawesi and includes the Sangiric languages.

Table 1. Typical Central Philippine vowel inventory

i	ɨ	u
(e)		(o)
	a	

Consonant inventories are also relatively simple and do not vary much across the area surveyed here. A typical inventory of phonemic consonants for the Central Philippine group is shown in Table 2. The tap ɾ can have several historical sources. Most typically, it is an intervocalic allophone of /d/.

We also find palatal obstruents at various stages of phonemicization, typically resulting from the combination of alveolars preceding /ij/, e.g. Tagalog 3SG.NOM /sija/ → [ʃ(j)a], 'there' /dijan/ → [dʒ(j)an], 'stomach' /tijan/ → [tʃ(j)an]. In a rarer development, Boholano has developed a voiced alveopalatal affricate from a historical palatal glide (i.e. PMP *y > dʒ).

Table 2. Typical Central Philippine consonant inventory

	labial	**alveolar**	**palatal**	**velar**	**laryngeal**
voiceless stop	p	t		k	ʔ
voiced stop	b	d		g	
nasal	m	n		ŋ	
fricative		s			h
lateral		l			

tap/trill		ɾ			
glide	w		j		

Unusual segments in CSP languages include the fortis/heavy stops of Maranao, described by Lobel and Riwarung (2009), and the aspirated stops of Subanen, described by Lobel and Hall (2010). As Lobel and Hall (2010:336-337) note, these form part of a larger set of unusual reflexes of consonant clusters in the languages of Mindanao and northern Borneo, a fact that they tentatively attribute to language contact.

2.2 Phonotactics

The canonical lexical root in Philippine languages is a disyllable with the following template: CV(C).CV(C). On one analysis, there are no true vowel-initial syllables in lexical roots (Zorc 1977:52). Roots that appear to be vowel-initial (and are treated as vowel-initial orthographically) begin with a glottal stop.[3] Relatedly, there is a general lack of vowel hiatus in most CSP languages, as vowel hiatus relies on the possibility of onsetless syllables.[4] Root initial glottal stops, whether they are underlying or epenthetic, surface predictably with prefixation, as in /mag-(ʔ)abut/ (AV-reach) → [magʔabot], rather than *[magabot]. On the most transparent analysis, all syllables in lexical roots begin with a consonant while codas are optional.

Monosyllabic lexical roots are both rare and a relatively recent innovation in Central Philippine languages, having entered through loans and various processes of reduction. In several languages of the Sulu archipelago, the deletion of intervocalic /l/ has created monosyllables with long vowels. Tboli shows another pattern of historically truncated monosyllabic roots, e.g. PMP *epat > *fat* ‘four’, PAN *kaen > *ken* ‘eat’.

Affixes do not have the same constraints as lexical roots; they are often monosyllabic and need not contain onsets. Onsetless affixes are typically provided with an onset either through epenthesis or infixation, the latter which only applies at the left edge of the base. When onsetless suffixes attach to stems that end in a vowel, either deletion or epenthesis avoids vowel hiatus. This latter process can be seen in Tagalog and Tagakaulo in (1a) and (b), respectively. The fricative /h/ is often used in this epenthetic capacity as it is not phonemic in root final position but glides also fulfill this role as in Tagakaulo.

(1)	a.	*bagu-hin*	b.	*bagu-wun*	
		new-PV		new-PV	(Burton 2018)

Infixes typically are of a VC shape but obtain an onset from the stem, as shown again for Tagalog and Tagkaulo in (2).

(2)	a.	*s<um>agot*	b.	*t<um>ubag*	
		<AV>answer		<AV>answer	(Burton 2018)

[3] Central Tagbanwa is apparently the only language in the CSP zone that is described as contrasting vowel initial syllables with glottal initial syllables (Scebold 2003:30).

[4] Words that are written with two vowels orthographically in languages like Tagalog, e.g. *bait* ‘goodness’, are pronounced with an intervening glottal stop, e.g. [baʔit]. Zorc (1977:54) mentions Cuyunon and certain dialects of Tausug as exceptional in allowing vowel hiatus.

Gemination is relatively rare in the CSP zone but is attested in Bagobo, Mansakan, Kagayanen and at least one dialect of Bikol (Blust 2013:229). Phonemic glottal stop arises from the historical change PMP *q > ʔ which took place widely throughout the Philippines but the synchronic distribution of the glottal stop varies by language and region. For instance, PMP *baqeRu 'new', reduced historically to a disyllable, yields (Naga) Bikol *baʔgo*, Cebuano *bagʔo* (with metathesis), and Tagalog *ba:go* (with deletion and compensatory lengthening). These changes follow a general pattern as Cebuano does not allow ʔC clusters and Standard Tagalog does not allow either Cʔ or ʔC clusters. Similar cases of metathesis are triggered by syncope when the resulting cluster is excluded by the general phonotactics of a language. Some of these clusters are universally absent in certain subgroups. For instance, Zorc (1977) cites **nm* as an unattested cluster in Bisayan roots and one that is actively avoided in forms that undergo syncope, as in /inum-an/ drink-LV which yields [imnan] with metathesis of the nasal consonants after deletion of medial /u/.

There appears to be a gradated loss of root final glottal stop from south to north. In southern CSP languages, glottal stop is highly salient phonetically and does not appear to undergo (synchronic) deletion. In many languages of the northern Philippines, root final glottal stop has been lost completely. In Tagalog, which lies on the border, word final glottal stop is less phonetically salient than in the south and it is often lost in phrase medial position, occasionally with compensatory lengthening. But even within a single subgroup, we find variation in the distribution of glottal stop. In the three members of the Danao languages, Maranao allows stem/word final glottal stop but Iranun and Maguindanao have both eliminated it in this position.

To summarize the status of the glottal stop in CSP languages: (i) there is only one language that possibly shows a contrast between V and ʔV at the beginning of roots (Central Tagbanwa); (ii) root internally, some language allow ʔC, others Cʔ, while others allow neither; (iii) most but not all CSP languages contrast root-final ʔ with root-final V.

In most CSP languages, glides pattern like any other consonant in the native vocabulary, but in some languages, glides can form consonant clusters at the syllable edge. For instance, in the Jolo dialect of Tausug, we find monosyllables such as *awn* EXIST and *lawŋ* 'inside'. In onset position, we find languages such as Maranao and Tagkaulo where the historical perfective infix *<in> has been reduced to a single glide <y>. In these languages, onset clusters with *y* as a second member are commonly derived through infixation. Tagalog shows a historical pattern of intervocalic *l* deletion which occasionally gives rise to similar clusters, e.g. PMP *bulan > Tag. *buwan ~ bwan*.

The Bilic languages of Southern Mindanao are exceptional with regard to the typically simple syllable margins of Philippine languages. Tboli allows for a large number of typologically rare onset clusters that violate the principle of sonority sequencing with regard to manner and voicing, e.g. /btaŋ/ 'fall', /tboli/ 'Tboli'.[5]

Complex tautosyllabic clusters have also entered CSP languages through Spanish and English borrowings, e.g. Tagalog *plato* 'plate', *preno* 'brake'. An illustrative example is seen in the Spanish loan *sombrero*, which enters Tagalog at a very early stage as *sambalilo*, fully adapted to native Tagalog phonotactics, and again at a later stage as *sombrero*, with the non-native *br* cluster and free distribution of mid-vowels, which were originally word final allophones of high vowels.

[5] While these can be broken up with a schwa, according to Awed et al. (2004), schwa insertion is optional. Whether this schwa should be analyzed as underlying or epenthetic has not been addressed in the literature.

All Philippine languages allow heterosyllabic clusters although each language exhibits its own constraints and tendencies. Interestingly, such clusters may be innovative and do not generally reconstruct to PMP (Blust 2013:62). The only clusters found at the PMP level as reconstructed in Blust & Trussel (ongoing) are either nasal+stop sequences, e.g. **simbuR* 'to sprinkle', or the result of reduplicated monosyllables, e.g. **taktak* 'to fall, of many things at once'. However, there are many apparently reconstructable lexemes in Philippine languages which contain clusters that do not fit into either of the above patterns. It should be noted that gradient phonotactic patterns have not been examined systematically for languages of the CSP area and present a rich area for further study.[6]

2.2 Phonological processes

The phonology of most CSP languages is relatively transparent in that surface forms do not differ substantially from what would be posited as underlying forms. Attested processes include palatalization, lenition, fortition, metathesis, and compensatory lengthening, exemplified below.

2.2.1 Lenition

Tapping, a type of lenition, takes place in Tagalog morpheme internally, between a prefix-stem boundary as well as between a word-enclitic boundary. Tapping does not occur in Tagalog between proclitics and their following hosts, as seen in (3), although other languages show tapping in these contexts, too, as shown in (4) for Matigsalug Manobo.

(3) a. /da:~datiŋ/ → [da:ɾatiŋ]
IPFV~arrive

b. /aku=din/ → [ako ɾin]
1SG.NOM=also

c. /maŋa=dagaʔ/ → [maŋa=dagaʔ], *[maŋa=ɾagaʔ]

(4) Matigsalug Manobo, tapping
/me=datuʔ/ → [me ɾatuʔ]
PL=chief (Wang et al. 2006:3)

Other types of lenition can be found in Binukid (aka Western Bukidnon Manobo) (Blust 2013:236), where it applies productively with affixation, e.g. *balәy* 'house', *bә-valәy* 'build a house', *guraŋ* 'old', *mә-ɣuraŋ* 'old person; old'. Deletion of intervocalic /l/ is also common across the area and was clearly a historical process in Tagalog, as well, although it was not carried out to completion.

2.2.2 Palatalization

A palatalization processes takes place in Tagalog with the alveolar obstruents /t/, /s/ and /d/ before /j/, as shown in (5a-b). A phonetically less natural palatalization process also takes place with the sequence /ts/, transforming it to [tʃ], as in (5c).

[6] Zorc (1977:53) notes the existence of phonotactic constraints in heterosyllabic clusters but laments the lack of data to address its nature. For Austronesian languages outside the CSP area, see Coetzee and Pater (2008) for Muna (Southeast Sulawesi) and Benton (1971) for Pangasinan (Northern Luzon) for examples.

(5) a. /sija/ → [ʃja] ~ [ʃa] 3sg.NOM b. /tijan/ → [tʃjan] ~ [tʃan]. stomach c. /at saka/ → [tʃaka] and then

In Central Tagbanwa, we find a similar but more circumscribed palatalization rule: /t/ → [tʃ] / __i. A similar pattern, although less advanced, is found in other Central Philippine languages like Cebuano. Despite allophonic rules that create palatal or alveopalatal segments, no CSP language has fully phonemicized a palatal series of obstruents.

2.2.3 Syncope and metathesis

The canonical Austronesian root is disyllabic and trisyllabic stems are reduced to disyllables through an active rule of syncope in many CSP languages, exemplified by Agutaynen in (6).

(6) Agutaynen, syncope
a. /balet-en/ → [balten] respond-PV
b. /b<in>etaŋ/ → [bintaŋ] <PFV>put (Quakenbush et al. 2010:41)

In rarer cases, syncope has been attested across clitic boundaries, as described by Lobel and Riwarung (2009, 2011) for Maranao clitics, such as *səka* 2SG.NOM and *səkano* 2SG.NOM, shown in (7).

(7) Maranao
[dɤ.ʔɤ.mɪs.ka.no.ma.ɪ.lay]
/daʔ ami səkano ma-ilay/
NEG 1Pl.EX.GEN 2PL.NOM PV.POT-see
'We didn't see you (pl.)' (Lobel and Riwarung 2011:41)

When syncope creates a cluster that is otherwise unattested, a phonological process typically repairs the output. In Agutaynen, a debuccalization process C → ʔ repairs certain clusters, as shown in (8), while in other cases, metathesis is employed, as in (9).

(8) Agutaynen, syncope + debuccalization
/te~teled/ → tetled → [teʔled]
PROG~enter (Quakenbush et al. 2010:42)

(9) Agutaynen, syncope + metathesis
a. /pa-belag/ → pablag → [palbag]
CAUS-separate

b. /pa-belet/ → pablet → [palbet]
CAUS-borrow (Quakenbush et al. 2010:41)

Neither syncope nor metathesis are productive in Tagalog but both processes are richly attested in allomorphy, as seen in (10) (see Blust 1971 for the complex interaction of metathesis and assimilation in this pattern).

(10) Tagalog, metathesis

a. /atip-an/ → atpan → [aptan]
roof-LV

b. /silid-an/ → sildan → [sidlan]
room-LV (Bloomfield 1917:391)

2.2.4 Vowel reduction and harmony

Vowel reduction is not common in Central Philippine languages but found in several languages of Mindanao and Sulu, e.g. Sindangan Subanen (Arms 1996:5), as well as Bornean languages south of the Philippines. Lobel and Riwarung (2009, 2011) describe a rare and intriguing case of harmony in Maranao where two complementary sets of vowels have developed, a "lax" set, [ɪ, ə, o, a], and a corresponding "tense" set, [i, ɨ, u, ɤ]. They show that the set of consonants they term "heavy", represented as /p', t', k', s', h/, obligatorily trigger the tense allophones of the following vowels. The voiced stops /b, d, g/ optionally trigger the tensing of the following vowel, and all other consonants condition the lax set. Because the heavy/light distinction on consonants plays an important role in the morphology, there are minimal pairs for every verb, as exemplified in (11). The "future" is signaled by the change of a light stem initial consonant to its heavy counterpart, and the consequent vowel harmony.

(11) Maranao

a.	[t̪a.ʔa.man]	b.	[t̪'ɤ.ʔɤ.man]	
	/taʔam-an/		/t'aʔam-an/	
	taste-LV		FUT/taste-LV	(Lobel and Riwarung 2011:40)

Central Tagbanwa shows a rightwards vowel harmony process with prefixes, as in (12). Unlike Maranao, this process does not affect lexical stems and is restricted to the change /a/→[u] immediately following a syllable bearing /u/.

(12) Central Tagbanwa

a.	[pupuŋaralan]	b.	[pugputabas]	
	/pu-paŋ-aral-an/		/pug-pa-tabas/	
	IPFV-DIST-study-LV		AV.IPFV-CAUS-prune	(Scebold 2003:35)

2.3 Morphophonology

2.3.1 Infixation

Two productive infixes inherited from PAn, *<*um*> ACTOR VOICE and *<*in*> PERFECTIVE/BEGUN ASPECT, continue to play an important role in CSP languages and Philippine languages more generally (Reid 1992). They are positioned after the first consonant of the stem, as shown in (13) for Tagalog. Historically, both of these infixes could co-occur as shown in Bikolano (14), although this is only found in a small number of living languages (Lobel 2004).

(13)a.	*k<in>uːha-Ø*	b.	*k<um>uːha*	(14)	*k<um><in>uːha*
	<BEG>take-PV		<AV>take		<AV><BEG>take
	'taken'		'take'		'consequently took'

Infixation is often externalized altogether in a process which turns *<*um*> into *mu-* and *<*in*> into *ni-*, as found in Cebuano. Reflexes of *<*in*> have also been reduced to a single segment in Danao languages (e.g. Maranao *t*<*i*>*abas* <PFV>cut), Tboli, Mansakan, and elsewhere in Mindanao.

Other minor infixes occur, as well, in a large number of CSP languages. For instance, in Bikolano and several Bisayan languages we find a plural infix <Vr>, whose vowel harmonizes with the first vowel of the stem. Another widespread <*aŋ*> infix marks a different type of plurality.

2.3.2 Reduplication

Philippine languages tend to make heavy use of various types of reduplication for a vast number of purposes. Tagalog has two types of CV reduplication, one with and one without vowel length, as well as foot reduplication. CV reduplication without vowel length is found in agentive nominalization, shown in (15a), intensive formation, and elsewhere. CV reduplication with vowel length, shown in (15b), is used chiefly for imperfective/progressive aspect.

(15)a. *mag-na~na:kaw*
AV-NMLZ-steal
'thief'

b. *mag-na:~na:kaw*
AV-IPFV-steal
'will steal'

Foot reduplication in many cases is indistinguishable from full reduplication of the root, as shown in (16a), as most roots are disyllabic. However, larger stems, as in (16b), demonstrate that no process of reduplication in Tagalog copies more than a foot.

(16)a. *ma-ganda~ganda=sila*
ADJ-MODER~beauty=3PL.NOM
'They are moderately beautiful.'

b. *bali:~bali:ta?*
MODER~news
'gossip'

Other languages, such as Central Tagbanwa, possess full word reduplication without such a maximality constraint, as seen in (17).

Central Tagbanwa

(17)a. *naka-tohod*
LOC-forest
'in the forest'

b. *naka-tohod~naka-tohod*
LOC-forest~LOC-forest
'deep in the forest' (Scebold 2003:42)

Multiple processes of reduplication can take place in the same word, as shown in Tagalog (18a), where (aspectual) CV reduplication applies to a stem that has already undergone (iterative) foot reduplication and in (18b), where (imperfective) CV: reduplication has applied to a stem that has undergone (intensive) CV reduplication.

(18)a. *mag-ha:~hanap~hanap*
AV-IPFV~ITER~search
'will keep searching'

b. *p*<*in*>*ag-sa:~sa~sabi*
<BEG>TR-IPFV~INTNS~search
'what is being said (intensively)'

Whereas Tagalog reduplication simply truncates a base that has more than two syllables, Cebuano and Bikol employ reduplication with fixed segmentalism for the same aim. Thus, for a trisyllabic Cebuano stem like *padala* 'send' we find *p<ulu>~padala*, where the first consonant of the stem has been copied and the following *ulu* is infixed, instead of **padala~padala* or **pada~padala* (see Mattes 2014:76 for additional complexities).

Word-based reduplication should be differentiated from a robustly syntactic process of reduplication which employs the linker or genitive case marking. These types of reduplication, shown for Central Tagbanwa in (19) and Tagalog in (20) (cf. Schachter and Otanes 1972:398), usually indicate repetitive action and are never affected by maximality constraints. Such constructions typically allow pronominal and other clitics to intervene between the base and the reduplicant, as in (20).

(19) Central Tagbanwa
t<um>umpok a t<um>umpok
<AV>pile LNK <AV>pile
'kept piling up' (Scebold 2003:57)

(20) *k<um>a:ʔin ako naŋ k<um>a:ʔin*
<AV.BEG>eat 1SG.NOM GEN <AV.BEG>eat
'I kept eating and eating.'

2.3.3 Nasal substitution

Languages of the CSP zone, like many other Malayo-Polynesian languages, display a morphophonological process termed "nasal substitution" with cognates of the sister prefixes PMP **paŋ-* DISTRIBUTIVE and **maŋ-* ACTOR VOICE + DISTRIBUTIVE (§3.4.2). Nasal substitution refers to assimilation of the final nasal of these prefixes to the place of articulation of the stem-initial consonant accompanied by deletion of the latter, as in Tagalog (21).[16]

(21) /maŋ-baril/ → [mamaril]
AV.DIST-gun
'shoot'

The deletion of the stem onset after nasal assimilation is not entirely predictable in Tagalog and other Central Philippine languages. Zuraw (2000) proposes a multifactorial analysis of this deletion for Tagalog, which must take into account the features of the first segment of the stem, as well as the stem's semantics and frequency. In other CSP languages, nasal substitution patterns are completely predictable on the basis of phonology alone, typically with stem-initial voiceless segments undergoing deletion and voiced segments being maintained (Blust 2004).

2.4 Stress and prosody

The vast of majority Philippine languages have a phonemic stress/prominence distinction on roots which has long posed a challenge for reconstruction. As discussed in Kaufman &

[16] The nasal coda of the prefixes that trigger nasal substitution are often represented by N, a placeless nasal with special morphophonological properties. Blust (2004) reviews nasal substitution patterns across Malayo-Polynesian languages.

Himmelmann (this volume), the basic feature that underlies the Philippine penultimate vs. final (aka paroxytone vs. oxytone) stress distinction is probably a vowel length contrast in the penultimate syllable. Central Philippine languages differ in whether closed penultimate syllables attract stress in the same way. In Tagalog, penultimate closed syllables do not attract pitch prominence nor can they co-occur with a long vowel and are thus predictably unaccented. In the Bisayan languages, on the other hand, closed penultimate syllables do attract pitch prominence on par with syllables containing a long vowel. Thus, a root like /dakdak/, in isolation, would surface as [dakˈdak] in Tagalog but [ˈdakdak] in Cebuano.

As noted by Blust (2013:251) and Kaufman & Himmelmann (this volume), prosody is not phonemic in several languages of the southern Philippines. Revel-Macdonald (1979:63) describes a general absence of phonemic accentual distinctions in Palawan but the presence of final syllable lengthening, which gives the impression of final stress. The lack of contrastive prosody (penultimate long vowels) appears to be a contact feature in this area. Pallesen (1985) observes that the Tausug of Sulu lacks the prosodic distinctions found in Central Philippine languages but that the Tausug of Palawan, which originated in 19th century Sulu, maintains the distinctions found in other Central Philippine languages, concluding that the loss of this distinction in Sulu is a relatively recent phenomenon that came about through contact with Sama languages, which show predictable penultimate word stress.

Other languages of the CSP zone without contrastive accent include Central Tagbanwa, which shows variable stress (Scebold 2003:27), Agutaynen, described by Quakenbush et al. (2010:40) as having penultimate phrase-based stress, Matigsalug Manobo, which shows regular penultimate word based stress (Wang et al. 2006:3), Maranao (Lobel and Riwarung 2011), and Tboli, which shows regular word final stress (Forsberg 1992).

CSP languages often employ vowel length, generally referred to as "contrastive stress" or "accent" in the literature, as a prosodic morpheme. Zorc (1977:64-67) discusses three types of morphological accent in the Bisayan languages which he takes to be part of the exponence of certain affixes. He notes, for instance, that in the Warayan subgroup of Bisayan, a prefix *ha-*, which derives adjectives indicating dimension and distance, co-occurs with penultimate stress. Thus, a root like *raˈyuʔ* 'distance' which shows final stress in isolation surfaces with penultimate stress with this prefix: *ha-ˈrayuʔ* 'far'. This apparent accent shift is likely due to the addition of vowel length to the penultimate syllable of the prefixed form (e.g. /ha-raːyuʔ/). Other Bisayan affixes co-occur with final stress and Zorc terms these "ultima-accent affixes", for instance, the prefix *manog-* 'on the verge of'. When attaching to a stem with penultimate stress like *ˈtapus* 'finish', the derived form *maˌnog-taˈpus* has final stress. Finally, Zorc discusses affixes that appear to flip the stress of the stem with final stress stems taking penultimate stress and vice versa.

The morphological use of vowel length and stress in the Central Philippine languages is still largely uncharted territory. Even for Tagalog, the best studied language of the CSP region, the facts remain elusive and not well understood. Little progress has been made since Zorc 1977 and some following work may have obscured these matters by ignoring the crucial role of vowel length in favor of a purely stress based analysis.

3 Morphology

The morphology of most Philippine languages is highly complex along several dimensions: (i) a large proportion of morphemes are multifunctional and take on distinct meanings in different

morphological contexts; (ii) much of the morphology is portmanteau, yielding a prototypical "fusional" language in Sapir's (1921) classic typology; (iii) the exponence of a morpheme, i.e., how a set of features are expressed on the surface, is often dependent on what other morphemes are present in the word. The discussion of root classes and lexical categories is handled in Kaufman (this volume). Below, aspect morphology (§3.1), voice morphology (§3.2), a variety of common derivational functions that typically fall under the heading of "mode" (§3.3), the causative (§3.4), and negation (§3.5).

3.1 Aspect

Although often described in terms of tense in the literature (e.g. McKaughan 1958, Wolff 1973, Zorc 1977 *inter alia*) the temporal inflections of Philippine languages uniformly indicate aspect rather than tense, with the possible exception of Iraya (Reid 2017). Voice and aspect are grammatically prominent and paradigmatically interconnected in most Philippine languages (cf. Reid 1992; Ross 2002; Himmelmann 2005). This can be seen in the Tagalog voice/aspect paradigm shown in Table 3, where the voice marker disappears unexpectedly in the prospective aspect of the actor voice paradigm and in the perfective of the patient voice paradigm.

Table 3. Fragment of the Tagalog voice aspect paradigm

baːsag 'break'	**Actor** *<um>*	**Patient** *-in*	**Locative** *-an*	**Conveyance** *i-*
Neutral	*b**<um>**aːsag*	*basaːg**-in***	*basaːg**-an***	***i-**baːsag*
Perfective	*b**<um>**aːsag*	*b<in>aːsag*	*b<in>aːsag**-an***	***i-**b<in>aːsag*
Progressive	*b**<um>**aː~baːsag*	*b<in>aː~baːsag*	*b<in>aː~baːsag**-an***	***i-**b<in>aː~baːsag*
Prospective	*baː~baːsag*	*baː~basaːg**-in***	*baː~basaːg**-an***	***i-**baː~baːsag*

A subset of Central Philippine languages display three primary aspects which can be termed perfective, progressive and prospective.[20] The three way distinction may arise from two atomic features corresponding to reflexes of *<in> and *CV reduplication, as in (22).

(22)

atomic features		compositional meanings	
<in>	BEGUN	<in>	**perfective**
CV~	IMPERFECTIVE	<in>CV~	**progressive**
		CV~	**prospective**

The feature combination [+begun, -imperfective] is interpreted as perfective, [+begun, +imperfective] as progressive, and [-begun, +imperfective] as prospective. Thus, while none of the surface aspects are indicated uniquely by a morpheme, they are derived in a compositional

[20] The prospective, which is used for unbegun action, is also referred to as "contemplated", "future" and "irrealis", all of which are, strictly speaking, inappropriate labels. "Contemplated" suggests cognition on the part of an agent; "future" designates a tense rather than an aspect; "irrealis" suggests that the form would be obligatory in negated and counterfactual contexts, although this is not the case.

manner (see Otanes 1966, De Guzman 1978 and Reid 1992 for different feature based approaches to this paradigm).[21]

Aspect marking is most often obligatorily on finite verbs although in some languages, such as Cebuano and Agutaynen, a single form will be used for the imperfective/prospective and the infinitive thus yielding a two-way distinction. Such languages can be said to conflate the historical unmarked and prospective aspects into a general 'unrealized' inflection (Reid 1992:74).

In addition to the major aspects shown in the above tables, most languages also possess minor aspects like Tagalog's recent perfective and immediate prospective, shown in (23).

(23)a. *ka-raː~ratiŋ ko laŋ*
RCT$_1$-RCT$_2$~arrive 1SG.GEN only
'I just arrived.'

b. *pa-ratiŋ na ako*
IMMD-arrive already 1SG.NOM
'I'm about to arrive.'

They are minor both in their frequency and in their emphatic interpretation, in contrast to the basic aspect categories. The syntax of the recent perfective is also distinct from the major aspects. In Tagalog and other Central Philippine languages, voice is neutralized and genitive case is assigned to what would normally be the nominative case marked argument. The recent perfective cannot be negated and may show additional syntactic restrictions, as well.

Other languages of the CSP area appear to have expanded this system more dramatically using the PMP mode prefix **paR*- as a durative (e.g. Aklanon, which Zorc 1977 analyzes as having six aspects). The use of a **paR*- reflex as a durative can also be seen in Cotabato Manobo (Kerr 1988:8), where *eg*- (< PMP **paR*-) indicates the progressive and CV reduplication no longer plays any role in the aspect paradigm (i.e. √-*en* PROSPECTIVE, *eg*-√-*en* PROGRESSIVE and <*in*>√ PERFECTIVE). The neighboring Danao languages also use a reflex of **paR*- (*pe*-) for what is signaled by reduplication in Tagalog, as seen in Table 4.

Table 4. Maranao voice aspect paradigm

	Actor	**Patient**	**Locative**	**Conveyance**
Neutral	*t<om>abas*	*tabas-en*	*tabas-an*	*i-tabas*
Perfective	*t<omi>abas*	*t<i>abas*	*t<i>abas-an*	*i-ni-tabas*
Progressive	*pe-tebas*	*pe-tebas-en*	*pe-tebas-an*	*i-pe-tebas*
Immediate prospective	*tebas*	*tebas-en*	*tebas-an*	*i-tebas*
Imperative	*tabas*	*tabas-a*	*tabas-i*	

In many languages, disyllabic reduplication indicates repetitive action and can be considered an aspectual category as well, although it is rarely included as part of the basic aspect paradigm in the descriptive literature and perhaps rightly so; unlike CV imperfective or

[21] PAn *CV~ may have originally marked the imperfective or durative while *<*in*> appears to have marked the perfective (Wolff 1973, Zorc 1977, Reid 1992, Ross 1995, 2002). Reid (1992) argues that *<*in*> innovatively spreads into the progressive in Central Philippine languages, where it comes to signal [+begun].

progressive reduplication, disyllabic repetitive reduplication is never seen to interact with mood, negation or voice.

The combination of *<*in*> with the composite actor voice markers beginning with *m-* (i.e. PMP **maŋ-* AV.DIST, **maki-* AV.SOC, **maR-* AV.MID, **maka-* AV.POT) typically yields *n-* initial forms without infixation (e.g. *naŋ-*, *naki-*, *nag-*, *naka-*). This "externalization" of *<in> postdates PMP, as we also find CSP languages that reflect **m<in>aR-* as *mig-* rather than *nag-*, showing that the full historical form was reduced in diverse ways after the break-up of the major Philippine subgroups.

In negated clauses, aspect is often indicated by the choice of negator and the verb is left unmarked or marked with an aspect neutral inflection. An example of this is seen in Sarangani Manobo, where aspect is marked on the verb in (24a-b) but through negation in (24c-d). Similar examples could be produced for most Bisayan languages, as well.

(24) Sarangani Manobo

a. *t<om>edogi se bayi*
<AV>sleep NOM woman
'The woman will sleep.'

b. *t<im>edogi se bayi*
<AV.PFV>sleep NOM woman
'The woman slept.'

c. *edek tedogi se bayi*
NEG sleep NOM woman
'The woman will not go to sleep.'

d. *weda? tedogi se bayi*
NEG.EXT sleep NOM woman
'The woman didn't sleep.' (Dubois 1976:20)

A more holistic understanding of aspect in Philippine-type languages must take into account both "inner aspect", i.e. perfective, progressive, prospective, as marked with bound morphology, together with "outer aspect", as marked by enclitics, typically descendants of PMP **=dena* 'already' (almost always reduced to a monosyllable) and **=pa* 'still'. Aspectual clitics in Philippine languages play a larger role than might be gleaned from their English glosses and are near obligatory in certain types of contexts. Outer aspect markers are both morphologically external to perfective, progressive and prospective morphology and also involve higher level pragmatics. Reflexes of PMP **=dena* 'already' place a situation before an *expected* time while PMP **=pa* 'still' places a situation after such a time.

3.2 Voice

Voice is a pivotal feature of the morphosyntax of all Philippine and Philippine-type languages.[23] The Philippine-type alignment system is generally understood to select a particular participant as the nominative argument (or absolutive, depending on the analysis) using one of several voice morphemes. This argument is typically interpreted definitely and can stand alone without an associated predicate. It is in some sense a privileged argument but its cross-linguistic status vis a vis subject and topic remains debated.

Agents of non-actor voice verbs are uniformly expressed in the genitive case in Philippine languages. Notional objects, when not selected by the voice morphology to become

[23] On the Bornean side, Lobel (2013:150) locates the southern border of the full voice system in the area of "Brunei Dusun, Kolod, Tingalan, Abai Sembuak/Tubu, Bulusu, and Tidung languages, although a handful of non-Philippine-type languages exist north of this hypothetical line." In Sulawesi, the full voice system seems to be continued only in the Mongondow-Gorontalo (or "Gorontalic") languages, as well as the Minahasan and Sangiric subgroups.

nominative arguments, are either expressed as genitives (as in Tagalog), as obliques (as in Cebuano), with something like a dedicated object marker (as in Maranao and Ivatan) or with the linker (as in the Bikol example below and more generally in Kapampangan).

The four primary voices are the actor voice, patient voice, locative voice and conveyance voice, as seen earlier in Table 5.[24] The exponence of these voice markers in CSP languages do not differ drastically from their PMP reconstructions.

Table 5. Common CSP reflexes of PMP voice markers

Voice	**PMP reconstruction**	**Common CSP reflexes**
actor voice	*<um>	<um>, m-, mu-
patient voice	*-en	-in, -un, -ɨn
locative voice	*-an	-an
conveyance voice	*(h)i-[25]	ʔi-, hi-, ∅

The basic use of the voice markers is very consistent across the CSP range with the exception of the Sama languages and, to a lesser extent, the Bilic languages. The system can be illustrated with the Naga Bikol examples in (25).

(25) Naga Bikol

a. *nag-bakal* *aku=ŋ* *bagas*
AV.BEG-buy 1SG.NOM=LNK rice
'I bought rice.'

b. *b<in>akal-∅* *ko* *an* *bagas*
<BEG>buy-PV 1SG.GEN NOM rice
'i bought the rice.'

c. *b<in>akal-an* *ko* *si* *hwan* *ki* *bagas*
<BEG>buy-LV 1SG.GEN NOM Juan OBL rice
'I bought some rice from Juan.'

d. *i-b<in>akal* *ko* *si* *hwan* *ki* *bagas*
CV-<BEG>buy 1SG.GEN NOM Juan OBL rice
'I bought some rice for Juan.' (McFarland 1974:104-105)

As can be seen, one participant is selected by the predicate to be the nominative argument while other participants are expressed in non-nominative cases. The actor voice selects the prototypical agent as the nominative argument; the patient voice typically selects an affected patient; the locative voice selects a locative, directional or other type of oblique argument as well as an

[24] What is termed here conveyance voice, following Wolff 1973, goes by several other names as well: circumstantial, instrumental, benefactive, secondary object, and theme voice, among others. See Blust (2002) and Ross (2002) for a review of the terminology and its history.

[25] The PMP cognate of the PAn conveyance voice marker **Si-* is predicted to be **hi-*, but this form only surfaces as such in Tausug and Samareño. Everywhere else, the initial *h* seems to have been eliminated in favor of a (possibly epenthetic) glottal stop. Nonetheless, because *h* is expected and these two languages were not in close contact with each other, the more common form *ʔi-* is thought to have come about through parallel innovation.

unaffected object; the conveyance voice selects a theme moving away from the agent as well as an instrumental or benefactee as the nominative argument.

The proper treatment of these voice markers remains an area of endless theorization and major debate in Austronesian linguistics. The earliest published analyses carried out by Spanish linguists and inherited by Bloomfield (1917) treated the patient, locative and conveyance voices as types of passive (e.g. direct passive, locative passive, etc.). It was recognized from the earliest point, however, that the putative "passives" of Philippine languages, which are fully transitive, were not equivalent to the Indo-European passive, a marked detransitive construction used primarily to background the agent. In the symmetrical analysis of Philippine-type voice (Foley 2008; Himmelmann 2005; Riesberg 2014), the system represents a unique type of alignment where all voices are equally marked, standing natural in opposition to accusative and ergative languages which typically display unmarked transitive and intransitive clauses. For the vast majority of CSP languages, it also holds true that there is no morphologically unmarked voice, just as in the Tagalog paradigm seen earlier. Proponents of an ergative analysis of the Philippine voice system (Starosta et al. 1982, De Guzman 1988, Gerdts 1988, Aldridge 2004, Liao 2004) argue that the actor voice appears less transitive than its non-actor voice counterparts. Although this is not the place to review the arguments for one analysis over another (but see Kaufman 2017), the principles of voice selection require basic explication.

There is widespread agreement that some type of referentiality largely determines voice selection (see Wolfenden 1961, Wolff 1966, Schachter 1976, McFarland 1978 for early treatments). Table 6, based on Tagalog but applicable more widely, abstracts away from many complications, additional factors, and cross-linguistic variation (Schachter 1976, Naylor 1986, Adams and Manaster-Ramer 1988, McFarland 1978, Latrouite 2011, Nolasco 2003) but captures the core basis for the alternation. When the agent is definite and the theme/patient is indefinite or absent, the predication will be expressed in the actor voice. When the theme/patient is definite, there is a strong tendency to employ the patient voice, regardless of the definiteness of the agent. With a verb of transfer and similar predicates, when the theme is indefinite but the recipient is definite, the locative voice will be selected. When a conveyed theme is definite, the conveyance voice will be selected, regardless of the definiteness of the agent and recipient.

Table 6. Voice selection in a typical CSP language

Agent	**Theme/Patient**	**Locative**	**Preferred Voice**
def	(indef)	–	ACTOR VOICE
def/indef	def	–	PATIENT VOICE
def/indef	(indef)	def	LOCATIVE VOICE
def/indef	def	def	CONVEYANCE VOICE

Definite referents can always be expressed as genitive agents and nominative arguments, while directional arguments are felicitously expressed in the oblique case regardless of their definiteness. What the pattern in Table 6 conspires to avoid is the expression of a definite undergoer as a non-nominative object.[27] If a previously introduced or otherwise familiar argument does surface as a non-nominative object, it typically receives a partitive interpretation or is understood to be less affected by the action (Nolasco 2003).

[27] This pattern holds throughout the CSP languages but some languages, such as Cebuano (Bell 1978), are argued to have a more flexible correspondence between the syntactic status of an argument and its definiteness.

In an intransitive predication with an indefinite subject, the subject is typically introduced with the use of an existential, as shown in (26a) (Schachter and Otanes 1972:279, but see Adams and Manaster-Ramer 1988 and Bell 1978 for additional wrinkles). The same holds for a bivalent predication in which neither argument has been previously introduced, as seen in (26b). This strategy is necessary to avoid the ordinarily definite interpretation of the nominative phrase.

(26)a. *may d<um>atiŋ*
EXT <AV.BEG>arrive
'Someone arrived.'

b. *may k<um>a:ʔin naŋ sa:giŋ*
EXT <AV.BEG>eat GEN banana
'Someone ate a banana.'

It should not be assumed that the patient voice is restricted to semantically bivalent predicates. Examples of the type in Tagalog (27) show that patient voice also selects affected subjects of monadic and even entity denoting predicates.

(27)a. *la:~laŋgam-in aŋ asu:kal*
IPFV~ant-PV NOM sugar
'The sugar will be "anted".'

b. *s<in>i:~sipon-Ø ako*
<BEG>IPFV~flu-PV 1SG.NOM
'I have the flu.' ('I'm being "flued".')

Similarly, the locative voice can select a recipient or location that we would consider part of the lexical semantics of the verb, as in (28), but it can just as easily "promote" an adjunct to become the nominative argument, as in (29).

(28) *b<in>igy-an ni rori naŋ pe:ra si pe:peŋ*
<BEG>give-LV GEN Rory GEN money NOM Pepeng
'Rory gave Pepeng money.'

(29) *in-iyak-an ni rori si pe:peŋ*
BEG-cry-LV GEN rory NOM Pepeng
'Rory cried to Pepeng.'

The locative voice can also alternate with the patient voice to indicate that the nominative argument is less affected by the action than would normally by assumed, as seen in the minimal pair in (30).

(30) a. *k<in>a:ʔin-Ø ni maria aŋ isdaʔ*
<BEG>eat-PV GEN Maria NOM fish
'Maria ate the fish.'

b. *k<in>aʔi:n-an ni maria aŋ isdaʔ*
<BEG>eat-LV GEN Maria NOM fish
'Maria ate from/at the fish.'

The conveyance voice (PAn *Si-) is difficult to characterize semantically in a unified manner. It selects benefactees, instrumentals and objects conveyed away from the agent as the nominative argument. These seemingly disparate functions can be disambiguated in a number of CSP languages with an emergent marker for each function, e.g. Tagalog *i-pag-* BENEFACTIVE, and *i-*

paŋ- INSTRUMENTAL, although the bare *i-* prefix in Tagalog is still as polysemous as its historical source.

In addition to the indicative/independent voice forms, there also exists a non-indicative/dependent paradigm. Wolff (1973:88) reconstructs this paradigm for the imperative and after certain "preverbs" while later work by Ross (2002) reconstructs it with a slightly wider range of functions. The CSP languages are crucial in understanding the role of the non-indicative forms in PMP, as they are preserved more faithfully here than in languages of the northern Philippines. In most MP languages outside the Philippines, the distinction between the indicative and non-indicative forms are also merged. In the northern and central Philippines, the paradigms are generally merged in favor of the indicative paradigm and are reduced in various ways south of the CSP zone.[28] Wherever the non-indicative paradigm is preserved, it is used in the imperative. This is seen in Batangas Tagalog (31) and Maranao (32). The dependent paradigm imperatives are distinguished from independent paradigm imperatives in most languages by the obligatory omission of a second person singular addressee pronoun, as in Batangas Tagalog, although there are rare exceptions to this, like Maranao.

Batangas Tagalog

(31) *buks-i (*mo) aŋ pintuʔan*
open-LV.DEP 2SG.GEN. NOM door
'Open the door!'

Maranao

(32) *tabas-a ŋka so dinis*
cut-PV.DEP 2SG.GEN NOM cloth
'Cut the cloth!' (McKaughan 1958:25)

In many Central Philippine languages, the dependent paradigm is also used in the negated perfective, as shown by Wolff (1973) for Samareño (33). This paradigm does not co-occur with imperfective reduplication or the perfective/begun *<in> infix (although they can occur in the recent perfective, see below).

Samareño (Waray)

(33)a. *waraʔ lakaw-Ø a baːtaʔ*
NEG.EXT go.away-AV.DEP NOM child
'The child did not go away.'

b. *waraʔ ku balik-a a sibiːsa*
NEG.EXT 1SG.GEN return-PV.DEP NOM beer
'I did not go back after the beer.'

c. *waraʔ ku hiŋalimt-i a isturya*

[28] In the majority of Austronesian languages, the independent locative voice *-an* survives with a nominalizer function and some remnant of *<um>* (typically melded with one of the mode prefixes as *m-*) survives in the actor voice. On the other hand, patient voice **-en* and conveyance voice **Si-* are widely lost as productive voice markers south of the Philippine languages, although the distinction may be carried out through different morphological means.

NEG.EXT 1SG.GEN forget-LV.DEP NOM story
'I did not forget the story.'

d. *waraʔ niya pilak-an an basuːra*
NEG.EXT 3SG.GEN throw.away-CV.DEP NOM garbage
'He did not throw the garbage away.' (Wolff 1973)

The dependent forms are also employed in temporal adjuncts, as seen in (34) and (35) (Stevens 1969, Zorc 1977) and the recent perfective (not shown here). These contexts are particularly interesting as the voice morphology selects a particular argument to promote, *kanya suwildu* in (34) and *baŋku* in (35), but no argument actually surfaces with nominative case.

Samareño (Waray)

(34) *pag-ta-tág-an=niya [sa kanya suwildu] [kanya nanay]...*
SBJV-ASP-give-CV.DEP=3SG.GEN OBL 3SG.GEN earning 3SG.GEN mother
'When he gives all of his earnings to his mother…' (Zorc 1977:139)

(35) *pag-liŋkur-i=niya han baŋku, na-rubaʔ*
SBJV-sit-LV.DEP=3SG.GEN GEN.DEF bench STA.PFV-break
'When he sat on the bench, it broke.' (Zorc 1977:139)

The Tboli voice system has been reshaped by the general loss of suffixes and case marking on full noun phrases. Here, there exists a general actor voice marked by *me-/* and a general undergoer voice marked by *ne-/<en>*, while the conveyance voice is left unmarked morphologically but still considered distinct. Tboli agent voice, undergoer voice and instrumental voice clauses are exemplified in (36).

Tboli

(36)a. *s<m>akay=le owoŋ yo ken ŋaʔ*
<AV>ride=3PL airplane that PL child
'The children rode in the airplane.'

b. *gel n-boʔ maʔ ɔu*
always UV-carry_on_back Father me
'Father always carried me on his back.'

c. *Ø-ɔfɔk Walan du asay*
CV-chop_down Walan it axe
'Walan chopped it down with an axe.' (Awed et al. 2004:79, 25)

We can also speak of composite minor voices, which appear to have been innovated more recently, often from combinations of inherited morphemes, and target adjuncts such as purposive clauses for promotion to pivot. The Tagalog prefix *ika-* (< PMP **(h)i-* CONVEYANCE VOICE + **ka-* STATIVE) and its cognate Sarangani Manobo exemplify this in (37) and (38).

(37) *ano aŋ ik<in>a-puːpunta niya duʔun?*
what NOM <BEG>REAS-IPFV~go 3SG.GEN there
'What's his reason for going there?'

Sarangani Manobo

(38) *yan se iŋke-opal ko*

that NOM REAS-anger 1SG.GEN
'That's why I became angry.' (Dubois 1976:67)

The CSP languages typically allow only one voice marker per word, but this is not the case in the languages of the northern Philippines. In many languages of North Luzon, reflexes of conveyance voice **(h)i-* combine with locative **-an* to form an unambiguous benefactive voice (Reid & Liao 2002:460). Such combinations are vanishingly rare in the CSP languages but may not be entirely absent, as seen in the Hiligaynon example in (39).[31]

Hiligaynon
(39) *i-lutuːʔ-an ko kamo sang paniʔudto*
CV-cook-LV 1SG.GEN 2PL.NOM GEN lunch
'I will cook lunch for you all.' (Wolfenden 1975:95)

3.3 Mode

There are several common verbal morphemes in CSP languages that are often treated under the somewhat vague header of "mode", a practice I continue here. These include the potentive (which subsumes both accidental and abilitative meanings), sociative and pluractional. Reflexives and reciprocals, as valency changing operations, are treated separately in §3.4.6.

3.3.1 Potentive

Nearly all CSP languages have a potentive paradigm, which is used to indicate both possible and unintentional action. This polysemy, which is remarkably stable across Austronesian languages, can be seen in the Tboli sentences in (40) and (41) with the *g(e)-* prefix, a reflex of PMP **ka-* (cf. Bennásar 1892:38-39 for the Tiruray cognate).

Tboli
(40) *nə g-tutuk kulu nib*
and POT-nail head Nib
'And Nib accidentally bumped his head.'

Tboli
(41) *g-uŋɔl-u udɛl sdoʔ fatu ləm law*
POT-hear-1SG.GEN voice pig across in cane
'I was able to hear the squeal of a pig in the cane across (the river).' (Forsberg 1992:92)

[31] Apparent combinations of voice markers do occur in the CSP languages when one voice marker derives the stem for the true voice marker. For instance, a Tagalog stem can be formed with locative nominalizer/voice marker *-an* and then go on to take the *mag-* actor voice prefix. Combinations of voice markers can also take on seemingly non-compositional functions, such as Tagalog *mag-tulug-tulug-an* AV-PRETEND~sleep-PRETEND 'to pretend to sleep', where both the reduplication and the *-an* suffix constitute multiple exponence of the 'pretendative'. But here there is no clear link between the pretendative function of *-an* and its more common locative voice function. Such cases are markedly different from *maŋ-i-* AV-CV- in Cordilleran languages, in which both the actor voice markers and the conveyance voice marker are playing a voice related role, the first determining the voice of the entire predicate and the second functioning as an applicative for objects moving away from the agent.

The potentive in CSP languages does not simply provide a way of emphasizing the accidental or unintentional nature of an action. It is obligatory in such contexts and as a corollary, the unmarked (non-potentive) form unambiguously denotes intentional action by an animate agent. This paradigm, which is contrasted with the unmarked "dynamic" voice paradigm in Table 7 for Tagalog, has a very distinct history involving the PAn prefix *_ka_-, whose original function may have involved possession (Kaufman 2011).

Table 7. The Tagalog potentive paradigm

	dynamic	**potentive**
ACTOR VOICE	*<um>*	*maka-*
PATIENT VOICE	*-in*	*ma-*
CONVEYANCE VOICE	*i-*	*ma-i-*
LOCATIVE VOICE	*-an*	*ma- -an*

The Tagalog potentive is transparently derived from the basic voice paradigm in the conveyance and locative voices with the addition of *ma-* but the actor and patient voices do not show clear correspondences. The potentive patient voice does not include a reflex of patient voice *-*en* and the potentive actor voice is not obviously related to other forms in the paradigm. This somewhat confusing picture, typical for Central Philippine languages, has a straightforward historical explanation. The *_ma_- prefix was originally a reduction of stative *_ka_- combined with actor voice *<*um*>, as a general *non-actor voice* potentive (Ross 1995:741). Historically, there was an opposition between an active clause such as (42a) and a passive-like stative clause, as in (47b), where the logical object would be the nominative argument. The latter is derived with the stative prefix *ka-* combined with the actor voice <*um*> followed by apheresis of the first syllable.

(42) a. *ʔ<um>u:bos*
<AV>finish
'to finish'

b. *k<um>a-ʔu:bos* → *ma-ʔu:bos*
<AV>STA-finish
'to get finished'

In all CSP languages that show a reflex of this *_ma_-, an agent can be introduced just as it is in a regular dynamic transitive clause yielding oppositions as in (43).

(43)a. *na-ʔu:bos ni bo:boy aŋ pagka:ʔin*
STA.BEG-finish GEN Boboy NOM food
'Boboy finished the food (accidentally)'

b. *<in>u:bos-Ø ni Bo:boy aŋ pag-ka:ʔin*
<BEG>-finish-PV GEN Boboy NOM food
'Boboy finished the food (purposefully)'

The use of the genitive in (43a) was most likely an innovation, and has apparently not taken place in many Cordilleran languages of the northern Philippines, which treat the paradigm

derived from **ma-* more along the lines of a passive than a transitive clause (cf. Reid & Liao 2004:464). The reanalysis of **ma-* from its original actor voice stative function to a potentive undergoer voice marker goes hand in hand with its appearance in other voices. The spread of **ma-* can be seen clearly in the comparison between Toratán (a Sangiric language of North Sulawesi, Himmelmann and Wolff 1999), Bikol Naga, and Tagalog (both Central Philippine), shown in Table 8. Note that *ma-* is labelled as a patient potentive in Table 8 due to sharing a case frame with the patient voice in CSP languages, as seen above in (48), but it is historically an actor voice form and is considered intransitive by Reid and Liao (2002:462).

Table 8. Potentive paradigms for three Philippine-type languages

	Toratán	**Bikol Naga**	**Tagalog**
Actor Voice	*maka-*	*maka-*	*maka-*
Patient Voice	***ma-***	***ma-***	***ma-***
Locative Voice	*ka- -an*	***ma-*** *-an*	***ma-*** *-an*
Conveyance Voice	*ka-*	*i-ka-*	***ma-****i-*

Toratán shows the most conservative paradigm, with *ka-* still used in both the conveyance and locative voices. It is innovative in having lost the *i-* in the potentive conveyance voice, but this is a recurring change seen to take place in Mindanao, as well. The *ma-* prefix has spread to the locative in Bikol Naga and additionally to the conveyance voice in Tagalog.[34]

The other oddity of the potentive paradigm is the actor voice counterpart to *ma-*, namely, *maka-*, which is derived from the combination of PMP **<um>* with the PMP causative **pa-* and the stative **ka-*. The original opposition between today's patient and actor voice potentive was thus not one of voice but one of causation.

3.3.2 Distributive

Many CSP languages express a distributive or pluractional meaning with a reflex of the PMP prefix **paŋ-* and its actor voice counterpart **maŋ-*. For certain predicates, this is obligatory. For instance, the act of fishing, by its nature, involves repeated action and does not have a single fish as its target. The use of the pluractional has thus become obligatory for forming the predicate 'to fish' in several CSP languages, including Tagalog. For other predicates, such as Tagalog *kuha* 'take', shown in (44), it is optional and adds a meaning ranging from repeated action, action on plural generic objects and unwanted persistence (De Guzman 1978).

(44) a. *k<um>uːha*
<AV>take
'to take'

b. *maŋ-[k]uːha*
AV.DIST-take
'to take (many)'

[34] The replacement of *ka-* with *ma-* in the locative and conveyance voices appears to have been a gradual and messy process in the Central Philippine languages. In many languages, including Tagalog, the conservative *ka- -an* and *i-ka-* coexist alongside the innovative *ma- -an* and *ma-i-* but are used with innovative meanings or with a limited set of roots.

Although the distributive most often occurs in the actor voice form with a cognate of **maŋ-*, it is not restricted to the actor voice. As exemplified by Tagalog (45) and Sarangani Manobo (46), the distributive can co-occur with any voice in most CSP languages.

(45) *iːlog na laːbis na p<in>aŋ-isdaʔ-an*
river LNK overly LNK <BEG>DIST-fish-LV
'an over-fished river'

Sarangani Manobo
(46) *i-m-pem-[b]egay dan se libro*
CV-PFV-DIST-give 3PL.GEN NOM book
'They gave out books.' (Dubois 1976:76)

South of the Philippines, the distributive takes on new functions, such as that of a dedicated anti-passive in certain South Sulawesi languages (Kaufman 2017), as well as the default marker of actor voice, as in Malayic languages.

3.3.3 Sociative

A morphological category found most commonly among Philippine languages is the so-called 'sociative', expressed with a reflex of PMP *paki- or its actor voice counterpart, *maki-. In most cases, this morpheme can be translated into English as 'with others', as in Tagalog (47), although this often does not capture the relation between the agent and the others.

(47) a. *maki-hiŋiʔ*
AV.SOC-request
'to request'

b. *maki-taːwa*
AV.SOC-laugh
'to laugh with others'

c. *maki-upoʔ*
AV.SOC-sit
'to sit with others'

The sociative often connotes copying the action of others for social purposes, a meaning which is more salient for some predicates, such as (47b), than for others. The predicate *makitaːwa* is typically interpreted as laughing because other people are laughing whereas the predicate *makiupoʔ* is simply to sit among others. The sociative need not denote a social activity in a positive sense. For instance, 'to fight' is often expressed with the sociative in CSP languages, e.g. Tagalog *maki-pag-aːway* (AV.SOC-TR-fight), Cebuano *makig-aːway* (AV.SOC-fight). The difference between the sociative mode versus the unmarked mode in such cases is subtle but the sociative appears to foreground an aspect of social exchange, even with predicates like 'fight'.[35]

3.3.4 Plural agent marking

It appears possible to reconstruct a PMP marker **si-* which necessitated a plural subject (reconstructed by Kitada 2019 as a sociative and by Liao 2011 as simultaneous aspect). In Central Philippine languages, we find a reflex in such forms as Tagalog *mag-si-takbo* (AV-PL-

[35] It appears that the imperative of the sociative, **paki-*, has developed in another direction, now signaling a polite request in a number of Philippine languages. Liao (2011) argues that there need not be a derivational relationship between **paki-* and **maki-* although the pragmatic link between the sociative function and polite requests is unlikely to be accidental.

run), where it serves to mark plurality. In the Bisayan languages, a reflex of this prefix indicates individuated action over a group, translated with 'each (subject)' (Zorc 1977:143).

Although it is rare for CSP languages to show obligatory number agreement with any argument, plural marking can be indicated simultaneously by several morphemes for emphasis, as in Tagalog (48), where the matrix clause predicate takes both the *si-* prefix as well as the *<aŋ>* infix, both independently indicating agent plurality. The subordinate verb again takes the plural marker *si-*, in addition to the pluractional marker *paŋ-*.

(48) *n<aŋ>ag-si-handa=ŋ* *mag-si-pam-[b]aril*
AV.BEG<PL>-PL-prepare=LNK AV.BEG-PL-DIST-shoot
'they prepared to go shooting' (Venago 1929:62)

Similarly, in Agutaynen (49), we find that the distributive **maŋ-* prefix has been reinterpreted as a plural agent prefix, which can co-occur with another plural marker *<Vr>*, commonly found in nearby Central Philippine languages, and the locative voice *-an* suffix used in its reciprocal function.

Agutaynen
(49) *mam-[p]ag-s<or>oay-an*
AV.PL-TR-<PL>fight-LV
'They will fight each other.' (Quakenbush et al. 2010:43)

Plural marking is often not uniform across word classes. In Tagalog, Agutaynen and elsewhere, adjectives with the uninflectable *ma-* prefix indicate plurality via CV-reduplication (without vowel length), e.g. Tagalog *ma-taba?* (ADJ-fat), *ma-ta~taba?* (ADJ-PL-fat). In Maranao, plurality on adjectives is marked with the *<aŋ>* infix, and in Cebuano, the *<g>* infix carries out the same function on dimension adjectives, e.g. *mu<g>bo?* (<PL>short), *da<g>ko?* (<PL>large).

3.3.5 Multifunctional *paR-/maR-

Reflexes of **paR-* (**maR-*, in the actor voice) can be found in almost all CSP languages although the range of functions associated with these morphemes differ from language to language. As Pittman (1966) first noted, Tagalog *mag-* has apparently contradictory functions, in some cases increasing valency, e.g. *<um>akyat* 'to ascend' vs. *mag-akyat* 'to bring something up', and in other cases, e.g. *<um>ahit* 'to shave others' vs. *mag-ahit* 'to shave one's self', decreasing valency. Kaufman (2009, 2018b) derives the apparently contradictory functions of this affix by viewing it as a historically complex combination of two components: the well attested causative prefix **pa-* and a middle voice prefix **R-*, which fused with the former. With some roots and paradigms, it is the causative *pa-* function which is meaningful while in other cases it is the middle voice whose interpretation prevails. The middle function of **R-* is also implicated in the durative, reciprocal and reflexive functions found with the **paR-/maR-* prefix. A typically mixed paradigm showing both the putative middle function and causative function of **paR-* is found in Palawano (Zorc 1971), shown in Table 9. Here, a reflex of **maR-* is found in the progressive of both intransitive and transitive actor voice paradigms but in other aspects it signals transitivity. Progressive aspect is often associated with decreased transitivity (Hopper & Thompson 1980) and thus appears to derive from middle voice **R-*. On the other hand, causative **pa-* is clearly responsible for the increased transitivity of the forms in the right hand column.

Table 9. Partial Palawano actor voice paradigm (Zorc 1971:70, with PMP etymologies added)

	Intransitive AV	**Transitive AV**
progressive	*məgC$_1$ə-* (*p<um>a-R-REDP~)	*məgC$_1$ə-* (*p<um>a-R-REDP~)
perfective	*<umin>* (*<um><in>)	*nəg-* (*p<um><in>a-R-)
unbegun/habitual	*<um>* (*<um>)	*məg-* (*p<um>a-R-)
participle	*pəg-* (*pa-R-)	*pəg-* (*pa-R-)

In many CSP languages south of Tagalog, the "plain" actor voice *<*um*> paradigm increasingly gives way to a **maR*- paradigm, as discussed by Liao (2004:106) and Lobel (2004, 2013:46-47). This prefix also appears to have been borrowed in several areas in the Philippines as the reflex of **R* often does not match regular sound correspondences (Liao 2004:107-12, Reid and Liao 2004:457).

3.3.6 Reciprocals and reflexives

There are two recurring strategies for forming reciprocals in CSP languages. The first, shown in Tagalog (50a), involves an apparent circumfix formally consisting of the actor voice prefix together with the locative nominalizer/voice suffix, i.e. **maR*-√-*an*, a formation which is also found in Malay (e.g. *bər-təŋkar-an* AV-fight-RECP). The second, exemplified by Samar-Leyte (50b), involves the **maR*- prefix together with the **ka*- prefix, one of whose functions is similar to English *co-*, deriving a partner in sharing something denoted by the stem. This later formation may only happen to overlap semantically with the reciprocal proper in (50a), as it more often refers specifically to two agents sharing in an activity.

	Tagalog		Samareño (Waray)
(50)a.	*nag-patay-an sila*	b.	*nag-ka-du:rug hira*
	AV-kill-RECP 3PL.NOM		AV-CO-sleep 3PL.NOM
	'They killed each other.'		'They slept together.' (Zorc 1977:144)

In some cases, the **maR*- prefix appears to express a reciprocal on its own, as in Tagalog *mag-kita?* AV-see 'to meet'. There are other reciprocal markers whose etymologies are not so clear. For instance, Tboli marks reciprocals with an *s-* prefix (likely derived from PMP **si-* discussed above), e.g. *tagak* 'to leave behind' *s-tagak* 'to leave each other'; *toboŋ* 'to help' *s-toboŋ* 'to help each other' (Forsberg 1992:91). In Binukid, as well as several Bisayan languages, the reciprocal is expressed with a circumfix whose first part is the *<in>* infix and the latter part is *-a?* or *-ay* (with *-ay* also appearing in the Bisayan languages), as seen in (51). Although both components of this circumfix occur in other derivations, they do not seem to be semantically related.

	Binukid
(51)	*m<i>g-b<in>ulig-a?*
	<PFV>AV.DUR-<RECP$_1$>help-RECP$_2$
	'They helped each other.' (Post and Gardner 1992:xxiv)

Reflexives are also commonly expressed with a descendant of **maR-* and stative reflexives are expressed with a reflex of PMP **maR-pa-ka-* AV.TR-CAUS-STA- in certain CSP languages among other areas (Blust 2003). Tagalog (52) exemplifies a remnant of this construction although it is not entirely productive as a reflexive.

(52)a. *mag-pa-ka-buːlag* AV.TR-CAUS-STA-blind 'make oneself blind'
b. *mag-pa-ka-matay* AV.TR-CAUS-STA-die 'kill oneself'
c. *mag-pa-ka-taːʔo* AV.TR-CAUS-STA-person 'be humane' ('make self a person')

3.3.7 Inchoative

The inchoative, termed by Zorc (1977:142) "essive", has barely been investigated from a comparative perspective. In many languages, it is signaled with a unique prefix, as shown in (53).

(53)a. Aklanon
nagiŋ-rayna si neli
AV.PFV.INCH-queen PL.NOM Neli
'Nellie became a queen.'
(Zorc 1977:142)

b. Tiruray
mente-eteu
AV.INCH-person
'to become a person'
(Bennásar 1892:40)

The inchoative form also allows for non-actor voice derivations. These were still current in the Tagalog of the early 20th century, as seen in (54), but are now obsolete.

(54) *aŋ maynila aŋ p<in>agin-pariːʔ-an niya*
NOM Manila NOM <BEG>INCH-priest-LV 3SG.GEN
'It was in Manila where he was ordained a priest.' (Lendoyro 1909:256)

Maranao uses a periphrastic construction, as in (55a), or a simple reflex of **maR-*, as in (55b), to express change of state.

(55)a. Maranao
mim-bɤloy a ator
AV.DIST-change LNK stone
'changed into a rock'

b. *m<iy>ag-ɤtor*
AV<PFV>-rock
'became a rock'

3.4 Causative

The PAn causative prefix **pa-* is perhaps the most stable affix in the entire PMP morphological inventory and is found in some form in all the CSP languages. The causative introduces a causer into the argument structure and can co-occur with any voice, mode and aspect. Abstracting away from various complications, Table 10 shows the canonical mapping of roles to arguments in a causative clause.

Table 10. Canonical role/case correspondences in the causative

	GEN	NOM	OBL
Actor voice	theme	causer	causee
Patient voice	causer	causee	theme
Conveyance voice	causer	theme	causee

In an actor voice causative clause, as in (56), the nominative argument is the causer while the theme is expressed just as an actor voice object would be expressed. The causee, on the other hand, is expressed as an oblique argument.

(56) *nag-pa-suːlat ako naŋ liːham sa estudyaːnte*
AV.BEG-CAUS-write 1SG.NOM GEN letter OBL student
'I had a student write a letter.'

In a patient voice causative clause, as in (57), it is always the causee that is selected as the nominative argument rather than the theme. The agent is assigned genitive case, as expected, and the theme, if expressed, is assigned genitive or objective case.

(57) *p<in>a-sulat-Ø ko naŋ liːham aŋ estudyante*
<BEG>CAUS-write-PV 1SG.GEN GEN letter NOM student
'I had the student write a letter.'

The conveyance voice consistently selects causative themes as the nominative argument, regardless of what voice is used to "promote" the notional object to nominative in a non-causative clause. The example in (58) shows how the causer is expressed as a genitive agent, as in the other non-actor voices, the causee is expressed as an oblique, and the theme or "notional object" becomes the nominative argument.

(58) *i-p<in>a-suːlat ko sa estudyaːnte aŋ liːham*
CV-<BEG>CAUS-write 1SG.GEN OBL student NOM letter
'I had the student write a letter.'

3.5 Negation

CSP languages are relatively rich in negators; distinct functional negators exist for perfective events, prospective events, prohibitives (imperatives), identification and existential predication. Few if any languages possess five distinct negators for each of these functions, but many languages show three and four-way distinctions. The negation inventories of five CSP languages are shown in Table 11.

Table 11. Negation in CSP languages

	Tagalog	Aklanon	N. Subanen	Maranao	Tboli

PERFECTIVE EVENT	*hindiʔ*	*ʔuwaʔ*	*ʔəndaʔ*	*diʔ*	*laʔ*
PROSPECTIVE EVENT	*hindiʔ*	*ʔindiʔ*	*ʔəndiʔ*	*diʔ*	*laʔ*
PROHIBITIVE	*huwag*	*ʔayaw*	*ʔəndiʔ*	*diʔ*	*béʔ*
IDENTIFICATION	*hindiʔ*	*bukon*	*gənnaʔ*	*kenaʔ*	*sundu*
EXISTENTIAL	*walaʔ*	*ʔuwaʔ*	*ʔəndaidun*	*daraʔ*	*(laʔ wən)*

What is termed here 'event' versus 'identification' negation is often framed in terms of lexical categories, e.g. verbal, nominal, and adjectival negation. Non-verbal negation can often be traced to a word meaning 'different'. For instance, Blust & Trussel (ongoing) reconstruct both PWMP **beken* 'negator of nominals, other, different' as well as PWMP **laqin* 'different', which also comes to function as a general negative marker in Sorsogon.

The distinction between perfective and prospective negation is uncommon, occurring mostly in Bisayan languages that employ the negative existential in perfective event-denoting contexts.

It is a common feature of Malayo-Polynesian languages outside the Philippines to combine the event negator with the existential to derive a negative existential (e.g. Malay *tidak ada* NEG EXT and *ti-ada* NEG-EXT), but most Philippine languages employ distinct unanalyzable roots for the existential and negative existential. As seen in Table 11, Tboli employs an analytic combination, as commonly found further south. There is a degree of fluidity between these functions, as shown by McFarland (1974:254-6). Nonetheless, there are several generalizations that can be made:

i. If a language has distinct negation for perfective events, it will be the same as the negative existential. (Subanen *ʔəndaidun*, above, exceptionally adds the formant *idun* in the negative existential.)
ii. If a language does not have a distinct prohibitive, this function will be carried out by the same form employed in the prospective.
iii. If a language does not have a distinct identification/non-verbal negator, this function will be carried out by the eventive/verbal negator.
iv. If a language does not have a distinct negative existential marker, this function will be carried out by the eventive negation in combination with the (positive) existential.

In a large number of CSP languages, certain negative contexts require the dependent verbal paradigm, as discussed by Wolff (1973) and Zorc (1977).

4. Elements of syntax

In this section, I present the basic word order across various phrase types (§4.1), and then take a closer look at word order within the noun phrase (§4.2) and the clause (§4.3). Finally, I look at the syntax of referential expressions: pronouns, demonstratives, case markers and the positioning of pronominal clitics (§4.4).

4.1 Basic word order relations

All the CSP languages are robustly head initial, as can be seen in the basic ordering relations exemplified by Tagalog in (59).

(59)

a. **Pred > Subj**
matali:no si bo:boy
smart NOM Boboy
'Boboy is smart'

b. **Noun > Possessor**
aŋ na:nay ni keŋkoy
NOM mother GEN Kengkoy
'Kengkoy's mother'

c. **Adj > Noun**
mataŋkad na baba:ʔe
tall LNK woman
'tall woman'

d. **Verb > Adv**
t<um>akbo naŋ mabilis
<AV>run GEN fast
'to run fast'

e. **Adposition > Noun**
ga:liŋ sa gu:bat
from OBL jungle
'from the jungle'

f. **Title > Name**
ginoʔo=ŋ reyes
mister=LNK Reyes
'Mister Reyes'

g. **Complementizer > Clause**
aka:laʔ ni dodoŋ na matali:no siya
thought GEN Dodong COMP smart 3SG.NOM
'Dodong thinks he's smart.'

h. **Noun > Relative Clause**
daga=ŋ p<in>atay-Ø ni=Kengkoy
rat=LNK <BEG>kill-PV GEN=kengkoy
'a rat killed by Kengkoy'

i. **Aux > Verb**
da:pat mag-madali=ka=na!
must AV.TR-hurry=2SG.NOM=already
'You should hurry up!'

j. **Comparative > Adjective > Standard**
lalo=ŋ mataŋkad sa kanya
more=LNK tall OBL 3SG.OBL
'taller than him/her'

k. **Negation > Verb**
hindiʔ s<um>ayaw
NEG <AV.BEG>dance
'didn't dance'

However, not all these relations are equal. Some, such as (e), (f), (g), (i), (j) and (k) are relatively strict or invariable. Others, such as (a), (b) and (d), allow for alternatives but with different semantic or pragmatic implications. A third category, which includes (c) and (h), represent tendencies but co-exist with equally unmarked alternative orders. We examine these in the following subsections.

4.2 Word order within the noun phrase

The vast majority of CSP languages are both head and dependent marking and possess a set of case marking determiners. While the order of case markers in relation to the noun phrase is strict, the order of certain modifiers within the noun phrase can be relatively flexible. The canonical order of elements in the Tagalog noun phrase is shown in (60). The elements in square brackets do not co-occur but rather represent two options for expressing possessors.

(66)

CASE	PRE-POSS	NUM	ADJ	ADJ	N	POST-POSS
aŋ	*[kanya=ŋ]*	*maŋa*	*ma-ga~ganda=ŋ*	*pula=ŋ*	*bulaklak*	*[niya]*
NOM	3SG.OBL=LNK	PL	ADJ-PL~beauty=LNK	red=LNK	flower	3SG.GEN

'his/her beautiful red flowers'

The case marker is in absolute initial position, as is the rule in Philippine languages, and this is followed by the position of the preposed possessor. The more common position for possessors is after the possessum, as shown on the right edge of the sequence although in rare cases, e.g. Hanunoo (Epo 2014), the preposed position appears to have become the norm. The preposed position only hosts pronominals in modern Tagalog although in earlier Tagalog, we find full NP possessors in this position, too, although stylistically marked. When possessors are preposed, they are always in the oblique case and never in the "pure" (typically *n-* initial) genitive case in CSP languages. Furthermore, they are typically connected to the following material in the phrase by the linker, as shown for Central Tagbanwa in (61a) (Scebold 2003:60), Tagalog in (61b), and Bikol in (61c).

	Central Tagbanwa		Tagalog		Naga Bikol	
(61)	a. *kanimi a bavoy*	b.	*inyo=ŋ ba:boy*	c.	*sa=indo=ŋ urig*	
	2PL.OBL LNK pig		2PL.OBL=LNK pig		OBL=2PL.OBL=LNK pig	
	'your (pl.) pig'		'your (pl.) pig'		'your (pl.) pig'	

In some languages, preposed oblique possessors have been described as inherently focused, as in Matigsalug Manobo (62). A better description for Tagalog would be that they are *focusable*, as opposed to the unfocusable enclitic pronominals.

	Matigsalug Manobo			
(62)	a. *ka anak ku*	b.	*ka keddì ne anak*	
	NOM child 1SG.GEN		NOM 1SG.OBL LNK child	
	'my child'		'my child (not his)'	(Wang et al. 2006:41)

Following this position we find the ubiquitous plural marker.[38] It is only the position of the case marker and the plural marker which are in a truly fixed position preceding the head noun. Following the plural marker, the canonical order of elements is adjective followed by noun, but this is variable in most CSP languages. In Tboli, where order appears to be more rigid, some adjectives must precede the noun, e.g. *tehe kimu* (former property), *dumu lan* (other path) (Forsberg 1992:39) but most follow the noun, e.g. *lan mahil* (path easy), *koyu lembaŋ* (tree large). For at least some adjectives, the position with regard to the noun is variable. As discussed by Donohue (2007:359-363), a rigid Noun-Adjective order emerges south of the CSP area and is common to languages of the Southeast Asian mainland. There is a marked difference between Central Philippine languages and those of the southern periphery in this regard, where the Bilic and Sama groups pattern similarly to languages of Indonesia.[39]

[38] Zorc (1977:103) claims that the plural marker (or "diversity marker") *maŋa* is found in all the Bisayan languages. Blust and Trussel (ongoing) reconstruct PMP **maŋa* as a prenominal plural marker. See Lynch et al. (2002: 90–91) for its history in Oceanic and Wu (2017) for a general look at plural markers in Austronesian, including the distribution of **maŋa*.

[39] Even languages of northern Sulawesi belonging to Blust's Philippine subgroup appear to show Donohue's (2007) southern pattern, e.g. Buol *botu moitomo* stone black (Zobel 2005:633). On the Bornean side, Kroeger (2005:411) describes the Kimaragang order of elements within the NP as: Determiner (Number) N (Possessor) (Modifier). It is only the unmarked position of the modifier that has shifted to the right edge when compared with the Central Philippine languages.

Demonstratives were left out of the template in (66) above because they are somewhat more difficult to generalize over in the CSP languages. Case is often marked syncretically on demonstratives, e.g. Tagalog *ito* 'this (neutral)', *nito* 'this (GEN)', *diːto* 'this/here (OBL)'. In Tagalog, a prenominal demonstrative takes the place of the case marker and is connected to the following material via the linker. It can also occur on the right edge of the noun phrase and here the argument is preceded by the expected case marker. Demonstratives can also sandwich the noun phrase for emphasis, as in (63).

(63) *ito=ŋ malaki=ŋ aːso=ŋ ito*
this=LNK big=LNK dog=LNK this
'this big dog'

In Northern Subanen (Daguman 2004:148), demonstratives are described as occurring only on the right edge of the noun phrase, but followed by relative clauses, as in (64).

Northern Subanen
(64) *s<in>aak-an su d-libun kətu nə mig-bələdyaʔ* ice cream
<RL>ask-LV NOM NM-woman that LNK AV.RL-sell ice cream
'...hey asked that lady who was selling ice-cream.' (Daguman 2004:159)

Obligatory classifiers are very rare in the CSP zone although they seem to exist in certain languages on the southern periphery. Daguman (2004:87) describes both sortal (e.g. *buuk* 'non-flat', *laad* 'flat', *tawan* 'human') and mensural (e.g. *dipa* 'arm span', *daŋaw* 'hand span') classifiers in Northern Subanen. These follow numeral modifiers and precede adjectives in the pre-head domain, as shown in (65).

Northern Subanen
(65) *...k=sala buuk g=əm-bagəl nə d=liun...*
NM=one CLF:non.flat NM=ADJ-big LNK NM=lion
'one big lion' (Daguman 2004:158)

The linker, which signals all types of modification, is common to the vast majority of Philippine languages but far rarer south of the CSP zone, even among Philippine-type languages and those of Blust's (2019) Philippine subgroup. The presence of the linker correlates with freer word order within the noun phrase. For instance, languages with linkers can typically place a relative clause before or after the phrase being modified. As the linker disappears towards the southern range of the CSP area, the order within the noun phrase becomes more rigid. The Bilic and Sama languages again pattern with their southern neighbors in lacking the linker and word order flexibility within the noun phrase (including the position of relative clauses in relation to their head noun).

4.3 Word order within the clause

As with all conservative MP languages, the CSP languages are almost without exception predicate initial across lexical category and clause type. Beyond the simple predicate-initial

generalization, the question of the basic order of phrases within the clause has never been answered definitively. Furthermore, as Himmelmann (2005:143) notes, there have been unwarranted claims of total freedom of phrasal order in the post-predicate domain. Nearly all CSP languages show the basic order shown in (66) for undergoer voice (i.e. non-actor voice) clauses and actor voice clauses.

(66) a. Undergoer voices
V AGEN PNOM

b. Actor voice
V (POBL/GEN) ANOM (POBL/GEN)

In the undergoer voices, there is a very strong tendency for the genitive marked agent to immediately follow the predicate head. In languages with impoverished case marking, this tendency becomes a rule. In the actor voice, the ordering relations appear to be less fixed although if there is an unmarked order, it tends to be one in which the nominative argument follows the patient.

The preverbal domain is typically reserved for pragmatically marked arguments and adjuncts (see Naylor 1975, Kroeger 1993, Kaufman 2005, Nagaya 2007 for Tagalog). All languages discussed here allow for topicalization of the nominative/absolutive argument to a preverbal position (Reid & Liao 2004:447). Typically, the fronted topic is followed by a dedicated topic marker, but in languages like Cebuano, there is topic fronting without a topic marker. In the unique case of Iraya, a language of northern Mindoro, most pronominal arguments must appear clause-initially, as exemplified in (67).

Iraya

(67)a. *Nay ʔinəm-en ʔag sapaʔ ŋuna*
1SG.GEN drink-PV DEF water now
'I'm drinking the water now.'

b. *kawu nay malyag.*
2SG.NOM 1SG.GEN like
'I like you.' (Reid 2017:34,27)

Zorc (1974) and Lobel (2013:188-193) also describe the shift to a pure actor voice, SVO syntax in main clauses in the Buhid language of southern Mindoro.

Oblique phrases, prepositional phrases and adjuncts can be topicalized in all the languages surveyed here. Genitive arguments and certain types of adjuncts cannot be topicalized so easily. The least extractable phrase is generally the actor voice object, which must occur post-verbally, as shown in (68). This restriction extends to relativization and cleft-like constructions in addition to topicalization. There is a sprawling theoretical literature on this pattern, which cannot be reviewed here (but see Kaufman 2017 for a summary). The constraint holds in much the same way across all CSP languages.

(68)a. *aŋ baːtaʔ ay k<um>aːʔin (naŋ) maŋga*
NOM child TOP <AV.BEG>eat GEN mango
'The child, ate the mango.'

b. **naŋ maŋga ay k<um>aːʔin aŋ baːtaʔ*
GEN mango TOP <AV.BEG>eat NOM child
(For, 'A mango, the child ate.')

The Central Philippine languages appear to have innovated a special focus position for fronted oblique arguments and adjuncts shown in (69).

(69) Foc[*sa mayni:la*]=*na*=*kami* *nag-a:~a:ral*
OBL Manila=already=1PL.EX.NOM AV.BEG-IPFV~study
'We already study *in Manila*.'

The focus fronted oblique phrase attracts second-position clitics and receives a cleft-like "exhaustive list" interpretation, i.e. 'It's in Manila (and nowhere else) that we study'. This construction is generally uncommon, if attested at all, in languages of the northern Philippines and most likely represents an innovation that took place in some subset of the CSP languages. In several Bisayan languages, focus fronting of an oblique phrase in this manner requires using the dependent paradigm of the verb.

4.4 Referential expressions

4.4.1 Pronouns

There are almost always distinct pronominal paradigms for the nominative, genitive and oblique cases. A typical example in this respect can be seen in the Maranao pronouns in Table 12 (McKaughan 1958, Kaufman 2010b). As is typical, clusivity is distinguished in the first person plural but relatively few languages have a distinct dual form, as Maranao does.

Table 12. Maranao pronoun paradigm

	Nom (bound)	**Nom (free)**	**Gen (bound)**	**Obl (free)**
1sg	*(a)ko*	*sakən*	*akən ~ ko*	*rakən*
1pl.ex	*kami*	*səkami*	*(a)mi*	*rəkami*
1+2 dual	*ta*	*səkta*	*ta*	*rəkta*
1pl.in	*tano*	*səktano*	*tano*	*rəktano*
2sg	*ka*	*səka*	*(ŋ)ka*	*rəka*
2pl	*kano*	*səkano*	*(n)iyo*	*rəkano*
3sg	*səkaniyan*	*səkaniyan*	*(n)iyan*	*rəkaniyan*
3pl	*siran*	*siran*	*(i)ran*	*kiran*

4.4.2 Demonstratives and deictics

Demonstratives and deictics in CSP languages typically distinguish three types of proximity: speaker proximate, hearer proximate and distal. Some languages, such as Matigsalug Manobo, shown in Table 13, distinguish four grades of proximity in deixis, although even in this language, the demonstratives only show the canonical three-way distinction.

Table 13. Matigsalug Manobo locative pronouns (Wang et al. 2006:28)

kayi, dini	here
due	there (within reach)
dutu	there (beyond reach but within sight)
diye?	way over there (nonspecific/out of sight)

Deictics are in most languages derived transparently from demonstratives with one of the PAn locative/directional markers *sa, *ka, *di (see Ross 2006 and Blust & Trussel ongoing for the reconstruction of these morphemes).

4.4.3 Case markers

Case markers have received ample attention from a historical perspective (Reid 2002, Blust 2005b, Reid 2007, Ross 2006, Blust 2015). I focus here on some salient features of typological interest. In the Central Philippine languages, case is often expressed syncretically with other referential and even temporal features. For instance, Waray employs three types of nominative and genitive case markers for full noun phrases: *ʔin* NOM indefinite, *ʔan* NOM past definite and *ʔit* NOM non-past definite, with genitive counterparts *hin, han, hit*, respectively (Zorc 1977:85). McFarland (1974) discusses similar specific/non-specific distinctions in the Legazpi Bikol case markers shown in (70) and (71). The (a) examples show that indefinite possessors and genitive agents are introduced by *ki* while definite ones are introduced by *kan.*

Legazpi Bikol

(70)a. *aruŋ ki lala:ki*
house GEN.INDEF man
'a man's house'

b. *aruŋ kan lala:ki*
house GEN.DEF man
'the man's house' (McFarland 1974:161)

Legazpi Bikol

(71)a. *pig-bakal ki lala:ki.*
PV.BEG-buy GEN.INDEF man
'bought by a man'

b. *pig-bakal kan lala:ki*
PV.BEG-buy GEN.DEF man
'bought by the man' (McFarland 1974:161)

Other varieties of Bikol make a subtle three-way distinction in referentiality, as seen for the Buhi dialect in Table 14. From the object marking in the examples in (72), we see that a generic object is marked by *nin*; a definite, but not yet "realized" object is marked by *nya*; and a definite, identifiable or "realized" object, is marked by *nyu.* As in Tagalog, the nominative phrase does not lend itself to an indefinite interpretation but still distinguishes what McFarland calls "definite" from "specific" arguments. In (72), because the action has not yet been realized, the subject receives the *a* marker. In (24a) and (c), because the action has been realized, the subject receives the specific *yu* marker.

Table 14. Buhi Bikol case markers (McFarland 1974:164)

	Nominative	**Genitive**	**Oblique**
Indefinite	–	*nin*	

Definite	*a*	*nya*	*sa*
Specific	*yu*	*nyu*	

(72)a. *aku* *yu* *nag-kaʔɨn* *nin* *aduːbu*
1SG.NOM NOM.SPEC AV.PFV-eat GEN.INDEF adobo
'I'm the one who ate adobo.'

b. *aku* *a* *nag-kaʔɨn* *nya* *aduːbu*
1SG.NOM NOM.DEF AV.PFV-eat GEN.DEF adobo
'I'm the one who will eat the adobo.'

c. *aku* *yu* *nag-kaʔɨn* *nyu* *aduːbu*
1SG.NOM NOM.SPEC AV.PFV-eat GEN.SPEC adobo
'I'm the one who ate the adobo.' (McFarland 1974:165)

In other languages which do not mark definiteness or specificity explicitly via case marking, the basic referentiality of an argument is largely predictable on the basis of grammatical function. There is some debate about whether these morphemes are inherently case markers or whether they have inherent referentiality related functions.[43]

In Table 15, we see case markers for common nouns (all nouns but personal names) in six CSP languages and in Table 16 we see their counterparts for personal names. It is immediately clear that Tboli diverges from the others in its reduced case system. All other languages make at least a three-way distinction between nominative case, genitive/ergative case, and an oblique case.

Table 15. Common noun case markers in six CSP languages

	Tagalog	**Aklanon**	**N. Subanen**	**Maranao**	**Tboli**
NOM/ABS	*aŋ*	*ro*	*su*	*so*	∅
GEN/ERG	*naŋ*	*it* (indef) *ku* (def)	*nə* *nu* (anaphoric)	*o*	∅
OBL	*sa*	*sa*	*sə* (local) *nə* (non-local)	*sa* (indef) *ko* (def)	*beʔ*

Table 16. Personal case markers in six CSP languages

	Tagalog	**Aklanon**	**N. Subanen**	**Maranao**	**Tboli**

[43] Himmelmann (2016) and Reid & Liao (2002:466) treat the Tagalog phrase marker *aŋ*, glossed NOMINATIVE here, as a definiteness marker of sorts without any inherent case features. Collins (2018), on the other hand, treats the same morpheme as a case marker without any inherent semantics at all. The fact that NP fragments with the nominative case marker always receive a referential interpretation (e.g. *dagaʔ!* 'a rat!' versus *aŋ dagaʔ!* 'the rat!') favors an analysis in which the case markers at least have some semantic features.

NOM/ABS	*si*	*si*	*si*	*si*	∅
GEN/ERG	*ni*	*ni*	*ni*	*i*	∅
OBL	*kay*	*kay*	*ni*	*ki*	∅

On the southern periphery of the CSP zone, as well as in Mindoro (Zorc 1974:577), word order takes becomes increasingly important in indicating grammatical relations. The example in (73) shows how actor voice objects and obliques may remain completely unmarked despite the existence of case markers in the language.

Cotabato Manobo

(73) *h<um>ated a sagiŋ kaut ta*
<AV>take 1SG.NOM banana kaut DET
'I will take some bananas to Kaut.' (Kerr 1988:13)

In Tboli, case is only distinguished on pronouns and the order of arguments in multi-argument clauses such as (74) is thus rigid.

Tboli

(74) *∅-oguh-en tum libun tum kun namak*
CV-hand.to-3SG.GEN that girl that 3SG.OBL betel.nut.quid
'He hands his own quid of betel nut to the girl.' (Forsberg 1992:78)

While the rich case marking system of Bikol languages shows that subtle referentiality distinctions can be made in the markers themselves, the basic definiteness distinction typical to Philippine type voice systems remains even in languages that have lost their case markers.

We can make the following generalizations about case marking in CSP languages:

i. There is a common three-way case system involving NOMINATIVE/ABSOLUTIVE, GENITIVE/ERGATIVE and OBLIQUE cases.
ii. The OBLIQUE case is employed for a wide range of directional/locative functions, as well as for marking definite objects of actor voice clauses, when this is allowed.
iii. The case of non-actor voice agents is always the same as that of possessors, hence labelled GENITIVE/ERGATIVE.
iv. Common noun phrases and personal names have distinct but morphologically related case markers.
v. Case marking is typically obligatory on all arguments.
vi. Case marking persists in accordance with the following hierarchy: pronouns > personal names > common nouns, such that it is lost first on common nouns and last on pronouns.

4.4.4 The positioning of clitics

All Philippine languages possess clitics whose positioning differs from full phrases. Pronominal and adverbial clitics are typically second position (aka Wackernagel) clitics following the first word and occasionally the first phrase within a clause-like syntactic domain (Kaufman 2010a). In languages like Maranao and Tagalog, pronominal arguments are in complementary distribution with full phrasal arguments, as seen in (75). When a potential clitic host precedes the predicate (in this case the progressive marker *diʔi*), a bound pronoun must typically attach to it, as shown in (76a), but this position is not available for full noun phrases, as shown in (75b).

Maranao

(75) a. *diʔi*[=*ako*] *ma-matiya*[*=*ako*] *sa kitab*
PROG=1SG.NOM AV-read=1SG.NOM OBL book
'I'm reading a book.'

b. *diʔi* [**so wataʔ*] *ma-matiya* [*so wataʔ*] *sa kitab.*
PROG NOM child AV-read NOM child OBL book
'The child is reading a book.' (Kaufman 2010b: 136)

In languages of this type, free pronouns are only used in predicate position, as independent fragments or as fronted topics. In other Central Philippine languages, such as Cebuano, long forms of the genitive and nominative pronouns show more syntactic freedom (Wolff 1966).

There are many cooccurrence constraints on pronominal clitics in the CSP languages with a fascinating variety of repair mechanisms and ordering patterns which cannot be discussed fully here (see Kaufman 2010a and references therein). The relative ordering of clitics is determined by up to three factors: prosody (shorter precedes longer), case (genitive precedes nominative) and person (1st person precedes 2nd person precedes 3rd person). Different constraints are active in different languages, but if a particular domain is active, it will always follow the above scales.

Clitic doubling, which is found occasionally in the northern Philippines, is rare in the CSP zone. Tboli, however, does show clitic doubling with certain preverbal elements, as seen in (76), where the second position clitic *le* doubles the nominative argument *kem dumu*.

Tboli

(76) *deŋ=le ma koyu kem dumu*
already=3PL.NOM AV.fetch wood PL companion
'The others already fetched some wood.' (Forsberg 1992:63)

5. Complex constructions

5.1 Finite complement clauses

All CSP languages allow for finite clause complements, as in (77). The embedded clause is generally introduced with the linker and has all the hallmarks of a main clause predicate.

(77) *s<in>aːbi-Ø ko sa iyo na gaː~gaw-in niya buːkas.*
<BEG>say-PV 1SG.GEN OBL 2SG LNK IPFV~do-PV 3SG.GEN tomorrow
'I told you that s/he will do (it) tomorrow.'

5.2 Questions and interrogative complements

When the interrogative phrase is a noun phrase, a cleft-like construction is required where the interrogative is in the predicate position and the remainder of the clause is embedded in a nominative phrase, as shown in (78).

(78) *ano aŋ s<in>a:bi-Ø niya?*
what NOM <BEG>say-PV 3SG.GEN
'What did s/he say?'

Interrogative complements are used in subordinate clauses as complements to matrix predicates of cognition as well as subjunctive type complements. These complements are typically identical to questions except that the interrogative phrase is introduced by a conditional marker, as in (79)-(81).

Central Tagbanwa
(79) *pog-tuʔma iŋ kali ka nag-giʔit.*
IPFV.AV-ask HYP where 2SG.NOM PFV.AV-depart
'He is asking where you came from.' (Scebold 2003:73)

Matigsalug Manobo
(80) *Su mig-inse sikandan ke hendei key eg-pa-bulus*
so AV.PFV-ask 3PL.NOM if where 1PL.EX.NOM AV.PROG-CAUS-continue
'So they asked where we were going.' (Wang et al. 2006:112)

Hanunoo
(81) *sabi-hun nimu sa kaŋku nu hayga*
tell-PV 2SG.GEN OBL 1SG.GEN COND why
'Tell me why (it's) that way.' (Epo 2014:22)

5.3 Nonfinite complement clauses

Clausal complementation with verbs of wanting, trying, and certain non-verbal predicates are typically non-finite and appear in a neutral form that does not indicate aspect, as shown in (82) and (83).

Cebuano
(82) *kinaháŋlan ni tibúʔ ŋa táwg-un aŋ páriʔ*
need GEN Tibo LNK call-PV NOM priest
'It is necessary for Tibo that a priest be called.' / 'Tibo needs to call a priest.'

Agutaynen
(83) *mambeŋ aŋ mag-pa-layog ta boradol*
fun LNK AV-CAUS-fly OBL kite
'It's fun to fly a kite.' (Quakenbush et al. 2010:13)

Note that voice marking is still present in most non-finite subordinate clauses. Other morphosyntactic categories discussed above, including the potentive, causative, reflexive, etc. can also appear in such contexts. In a small number of CSP languages, including Agutaynen, as

seen in (84), aspect in the subordinate clause agrees with the matrix predicate in what are typically non-finite contexts for other CSP languages.

Agutaynen

(84) a. *nam-[p]ag-t<ar>abaŋ-an tanira=ŋ naŋ-ayeg*
AV.PFV-PL-TR-<PL>help-LV 3PL.NOM=LNK AV.PFV.DIST-harvest
'They helped one another to harvest.'

b. *mam-[p]ag-t<ar>abaŋ-an tanira=ŋ maŋ-ayeg*
PFV.PL-TR-<PL>help-LV 3PL.NOM=LNK AV.DIST-harvest
'They will help one another to harvest.' (Quakenbush et al. 2010:20)

A less common type of complementation pattern attested in Central Philippine languages involves treating the subordinate predicate as a case marked complement, as shown in (85)-(86).

(85) *b<in>ilis-an ko aŋ pag-kaːʔin*
<BEG>fast-LV 1SG.GEN NOM GER-eat
'I sped up my eating.'

Cebuano

(86) *nag-siːge ug sunod sa iya=ŋ bukog*
AV.BEG-continue OBJ follow OBL 3SG.GEN=LNK bone
'He continues following his bones.'

5.3.1 Control patterns

Control refers to coreference between an argument in a matrix clause and a missing argument in a (typically non-finite) subordinate clause. Most CSP languages pattern as in (87), where an embedded agent co-referring with a matrix argument must be null.

(87) *gusto ko=ŋ tawaːg-an (*ko) si boboy*
want 1SG.GEN=LNK call-LV 1SG.GEN NOM Boboy
'I want to call Boboy.'

Conversely, the agent of a subordinate non-finite clause must be overt when it does not co-refer with a matrix clause argument, as in (88).

(88) *gusto ko=ŋ tawaːg-an mo ako*
want 1SG.GEN=LNK call-LV 2SG.GEN 1SG.NOM
'I want you to call me.'

As has been noted (Cena 1977, Kroeger 1993, Schachter 1976), the volitionality of the subordinate predicate determines which argument can be controlled, as seen in the minimal pair in (89).

(89)a. *gusto ko=ŋ tawaːg-an*
want 1SG.GEN=LNK call-LV

b. *gusto ko=ŋ ma-tawaːg-an*
want 1SG.GEN=LNK STA-call-LV

‘I want to call (someone).’ ‘I want to be called.’

This seems to hold true for at least the Central Philippine subgroup although this type of data is generally lacking for other subgroups.

5.3.2 The actor voice restriction

A more unusual phenomenon whose presence in Philippine languages has not received any notice is found in the Danao languages. For fully biclausal sentences, Maranao and Maguindanao show structures similar to Tagalog and other Central Philippine languages, as seen in Maguindanao (90), where the embedded verb is an infinitive in the locative voice.

Maguindanao

(90) *kalinian=neŋka tawag-an=ko seka?*
want=2SG.GEN call-LV=1SG.GEN 2SG.NOM
‘Do you want me to call you?’

However, the Danao languages have a reduced complement clause structure that requires the subordinate verb to appear in the actor voice, as in Maranao (91). Here, the undergoer of ‘call’ appears to obtain case from the matrix verb and is positioned in the matrix clause.

Maranao

(91) *t<in>ekaw-an ko seka tawag!*
<PFV>try-LV 1SG.GEN 2SG.NOM <AV>call
‘I tried to call you!’

This corresponds to the so called “actor voice constraint” discussed by Aldridge (2004) and Chang (2017) for Formosan languages and Kroeger (2014) for Kimaragang, a Dusunic language of Sabah, whereby certain types of subordinate clauses must be in the actor voice.

5.4 Adjunct clauses

Temporal adjuncts are most often formed via nominalization in CSP languages (Kaufman 2011). A typical structure is shown in Sarangani Manobo (92), which displays the combination of the gerundive *peg-* with the lack of a nominative case on either of the arguments.

Sarangani Manobo

(92) *peg-dineg te amay din kenyan*
GER-hear GEN father 3SG.GEN that.OBL
‘When his father heard that.’ (DuBois 1976:94)

6. Conclusion

This chapter has attempted to give a broad overview of the phonology, morphology and syntax of the CSP languages while focusing on several phenomena of interest that are characteristic of the region. I have also attempted to highlight areas in need of further research. In the phonology, gradient phonotactic generalizations have largely gone unexplored beyond Tagalog and the study of word prosody and intonation is also a rich and relatively untouched area. The CSP languages

have played a large role in our understanding of PMP morphosyntax but we still have an incomplete understanding of how the dependent paradigm was deployed as well as various types of subordination. The actor voice constraint has been presented here for the first time as a Philippine phenomenon, in addition to its presence in Sabah and Formosan languages.

Finally, a note on the general typology of the region. Himmelmann (2005) defines Philippine-type languages as having symmetric voice in addition to the following three characteristics:

(a) at least two formally and semantically different undergoer voices
(b) at least one non-local phrase marking clitic for nominal expressions
(c) pronominal second position clitics

These features, all of which are understood to be retentions from PMP, have eroded to various extents in the Bilic languages, the Sama languages (Kaufman this volume), and several languages of Mindoro, thus opening a typological rift within the CSP region. Thus, while the core of the CSP region, represented by familiar Central Philippine languages such as Tagalog and Cebuano, is typologically homogeneous, the outliers present fascinating departures from the norm.

Despite progress, there is still much work to be done in the description of CSP languages outside the Central Philippine group. While contact relations have been studied in Mindanao (see references in §1.1), no major effort has been made for other areas within the region. The Bilic languages, in addition to the languages of Palawan and Mindoro are in special need of further work with an eye towards contact relations. Blust (1992), examining the Tiruray lexicon, has already shown that the emerging picture is complex and multilayered. As Blust (2019) presents a new argument for the unity of a Philippine subgroup based on shared lexical innovations, it becomes even more urgent to understand the dual roles of contact and inheritance in the historical formation of Philippine languages.

References

Adams, Karen L. and Alexis Manaster-Ramer (1988). 'Some questions of topic/focus choice in Tagalog', *Oceanic Linguistics* 27(1/2): 79–101.

Aldridge, Edith (2004). 'Ergativity and Word Order in Austronesian Languages', PhD thesis, Cornell University.

Arms, D. G. (1996). 'Categories of the Sindangan Subanen verb', in Hein Steinhauer (ed.), *Papers in Austronesian Linguistics No. 3*, Canberra: Pacific Linguistics, 1-32.

Awed, Silin A., Lillian Underwood, and Vivian M. van Wynen (2004). *Tboli-English dictionary*. Manila: SIL.

Bell, Sarah J. (1978). 'Two Differences in Definiteness in Cebuano and Tagalog', *Oceanic Linguistics* 17(1): 1-9.

Bennásar, Guillermo (1892). *Observaciones gramaticales sobre la lengua Tiruray*. Manila: Imp. y litog. de M. Pérez.

Benton, R.A. (1971). 'Phonotactics of Pangasinan', PhD thesis, Department of Linguistics, University of Hawaii, Mānoa.

Bloomfield, Leonard. (1917). *Tagalog texts with grammatical analysis*. Urbana, Illinois: University of Illinois.
Blust, Robert (1971). 'A Tagalog consonant cluster conspiracy', *Philippine Journal of Linguistics* 2: 85-91.
Blust, Robert (1991). 'The Greater Central Philippines hypothesis', *Oceanic Linguistics* 30: 73-129.
Blust, Robert (1992). 'On speech strata in Tiruray', in Malcolm Ross (ed.), *Papers in Austronesian Linguistics, No. 1*. Canberra: Pacific Linguistics, 1-52.
Blust, Robert (2002). 'Notes on the history of focus in Austronesian languages', in Fay Wouk and Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics, 63–80.
Blust, Robert (2003). 'Three notes on Early Austronesian morphology', *Oceanic Linguistics* 42: 438-478.
Blust, Robert (2004). 'Austronesian nasal substitution: a survey', *Oceanic Linguistics* 43:73–148.
Blust, Robert (2005). 'A note on the history of genitive marking in Austronesian languages', *Oceanic Linguistics* 44: 215-22.
Blust, Robert A. (2013). *The Austronesian Languages*, revised edition. Canberra: Pacific Linguistics.
Blust, Robert (2015). 'The case-markers of Proto-Austronesian', *Oceanic Linguistics*. 54(2): 436-491.
Blust, Robert (2019). 'The resurrection of Proto-Philippines', *Oceanic Linguistics* 58(2): 153-256.
Blust, Robert and Steve Trussel. (ongoing). Austronesian Comparative Dictionary. URL: https://www.trussel2.com/acd/
Burton, Scott L. (2018). *Tagakaulo-English Dictionary*. Manila: SIL Philippines.
Burton, Scott L. (1996). 'A case study of lexical borrowing between language families: the East Mindanao and Manobo languages', PhD thesis, University of Texas at Arlington.
Cena, Resty M. (1977). 'Patient primacy in Tagalog', Paper presented at the LSA Annual Meeting, Chicago.
Chang, Henry Y. (2017). 'The AV-only restriction and locality in Formosan languages', *Tsing Hua Journal of Chinese Studies* 47(2): 231-254.
Coetzee, Andries W. and Pater, Joe (2008). 'Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic', *Natural Language and Linguistics Theory* 26: 289–337.
Collins, James N. (2018). 'Definiteness determined by syntax: A case study in Tagalog', *Natural Language and Linguistic Theory*.
Daguman, Josephina S. (2004). 'A grammar of Northern Subanen', PhD thesis, La Trobe University.
De Guzman, Videa P. (1988). 'Ergative analysis for Philippine languages: An analysis', in R. McGinn (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for Southeast Asia Studies, 323-345.
De Guzman, Videa P. (1978). 'Syntactic derivation of Tagalog verbs', *Oceanic Linguistics Special Publication No. 16*. Honolulu: University of Hawaii Press.
Donohue, Mark (2007). 'Word order in Austronesian from north to south and west to east', *Linguistic Typology* 11: 349–391.

Dubois, Carl .D. (1976). *Sarangani Manobo: an introductory guide*. Manila: Linguistic Society of the Philippines.
Elkins, Richard (1986). *Manobo-English Dictionary*. Honolulu: University of Hawai'i Press.
Epo, Yrrah Jane S. (2014). 'Discourse analysis of suyot: a Hanunuo-Mangyan folk narrative.' MA thesis. Payap University.
Fleischman, Eric (1981). 'The Danao Languages: Magindanaon, Iranun, Maranao, and Illanun', *Philippine Journal of Linguistics* 12: 57-78.
Foley, William A. (2008). 'The place of Philippine languages in a typology of voice systems', in Peter K. Austin and Simon Musgrave (eds.), *Voice and Grammatical Relation in Austronesian Languages*. Stanford: CSLI Publications, 22-44.
Forsberg, Vivian M. (1992). 'A Pedagogical Grammar of Tboli', *Studies in Philippine Lingiustics* 9: 1-110.
Gallman, II, Andrew Franklin. (1983). 'Proto East Mindanao and Its Internal Relationships', PhD thesis, University of Texas at Arlington.
Gerdts, D. B. (1988). 'Antipassives and causatives in Ilokano: Evidence for an ergative analysis', in R. McGinn (ed.), *Studies in Austronesian Linguistics*. Athens, OH: Ohio University Center for Southeast Asia Studies, 295-321.
Himmelmann, Nikolaus P. (1991). 'The Philippine challenge to Universal Grammar', Arbeitspapier Nr. 15. Köln: Institut für Sprachwissenschaft, Universität zu Köln.
Himmelmann, Nikolaus P. and John Ulrich Wolff. (1999). *Toratán (Ratahan)*. Languages of the World/Materials 130. Munich: Lincom Europa.
Himmelmann, Nikolaus P. (2005). 'The Austronesian languages of Asia and Madagascar: Typological characteristics', in K. Alexander Adelaar and Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*. London: Routledge, 110-181.
Himmelmann, Nikolaus P. (2016). 'Notes on 'noun phrase structure' in Tagalog', in Jens Fleischhauer, Anja Latrouite, and Rainer Osswald (eds.), *Explorations of the Syntax-Semantics Interface*. Düsseldorf: Dup, 319-341
Kaufman, Daniel and Nikolaus P. Himmelmann (this volume). 'Supersegmental Phonology'.
Kaufman, Daniel (2005). 'Aspects of pragmatic focus in Tagalog', in I Wayan Arka and Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies*. Canberra: Pacific Linguistics, 175–196.
Kaufman, Daniel (2010a). 'The morphosyntax of Tagalog clitics: A typological approach', PhD thesis, Cornell University.
Kaufman, Daniel (2010b). 'The grammar of clitics in Maranao', in Loren Billings and Nellke Goudswaard (eds.), *Piakandatu ami Dr. Howard P. McKaughan*. Manila: Linguistic Society of the Philippines and SIL, 132-157.
Kaufman, Daniel (2011). 'Exclamatives and temporal nominalizations in Austronesian', in Foong Ha Yap, Karen Grunow-Hårsta, and Janick Wrona (eds.), *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam: John Benjamins, 721-754.
Kaufman, Daniel (2017). 'Lexical category and alignment in Austronesian', in Lisa Travis, Jessica Coon, and Diane Massam (eds.), *Oxford Handbook of Ergativity*. Oxford: Oxford University Press.
Kaufman, Daniel (2018a). 'Austronesian predication and the emergence of biclausal clefts in Indonesian languages', in Sonja Riesberg, Asako Shiohara, and Atsuko Utsumi (eds.),

Perspectives on information structure in Austronesian languages. Berlin: Language Science Press, 207-245.
Kaufman, Daniel (2018b). 'Inner and outer causatives in Austronesian: A diachronic perspective'. *McGill Working Papers in Linguistics* 25(1): 201–215.
Kaufman, Daniel (this volume). 'Lexical categories'.
Kaufman, Daniel (this volume). 'The Sama-Bajaw languages'.
Kerr, Harland (1965). 'The case-marking and classifying function of Cotabato Manobo voice affixes', *Oceanic Linguistics* 4: 15–47.
Kerr, Harland (1988). 'Cotabato Manobo Grammar', *Studies in Philippine Linguistics* 7: 1–123.
Kroeger, Paul R. (1993). *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
Kroeger, Paul (2005). 'Kimaragang', in Alexander Adelaar and Nikolaus Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London & New York: Routledge, 397-428.
Kroeger, Paul (2014). 'An affectedness constraint in Kimaragang restructuring' in Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG14 Conference*. Stanford, CA: CSLI Publications, 283–303. http://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/papers/lfg14kroeger.pd
Latrouite, Anja (2011). 'Voice and Case in Tagalog: The coding of prominence and orientation', PhD thesis, University of Dusseldorf.
Lendoyro, Constantino (1909). *The Tagalog Language*. Manila: Juan Fajardo.
Liao, Hsiu-chuan (2004). 'Transitivity and ergativity in Formosan and Philippine languages', PhD thesis, Department of Linguistics, University of Hawai'i at Mānoa.
Liao, Hsiu-chuan (2011). 'Some morphosyntactic differences between Formosan and Philippine languages', *Language and Linguistics* 12: 845-876.
Lobel, Jason W. (2004). 'Old Bikol -um- vs. mag- and the loss of a morphological paradigm', *Oceanic Linguistics* 43: 469-497.
Lobel, Jason W. (2013). 'Philippine and North Bornean languages: issues in description, subgrouping, and reconstruction', PhD thesis, Department of Linguistics, University of Hawai'i at Mānoa.
Lobel, Jason W. and William C. Hall (2010). 'Southern Subanen Aspiration', *Oceanic Linguistics* 49(2):319–338.
Lobel, Jason W. and Labi Hadji Sarip Riwarung (2009). 'Maranao revisited: An overlooked consonant contrast and its implications for lexicography and grammar', *Oceanic Linguistics* 48(2):403–438.
Lobel, Jason W. and Labi Hadji Sarip Riwarung (2011). 'Maranao: a preliminary phonological sketch with supporting audio', *Language Documentation and Conservation* 5: 31-59. Available online at http://hdl.handle.net/10125/4487/.
Lobel, Jason W. (2004). 'Old Bikol -um- and mag- and the loss of a morphological paradigm.' *Oceanic Linguistics* 43(2): 469–497.
Lynch, John, Malcolm Ross, and Terry Crowley (2002). *The Oceanic Languages*. London/New York: Routledge.
McFarland, Curtis D. (1974). 'The Dialects of the Bikol Area', PhD thesis, Yale University.

McFarland, Curtis D. (1978). 'Definite objects and subject selection in Philippine languages', in Casilda Edrial-Luzares and Austin Hale (eds.), *Studies in Philippine Linguistics, Vol. 2*. Manila: Linguistic Society of the Philippines, 139–182.
McKaughan, Howard P. (1958). *The inflection and syntax of Maranao verbs*. Manila: Institute of National Language.
Nagaya, Naonori (2007). 'Information Structure and Constituent Order in Tagalog'. *Language and Linguistics* 8(1): 343–372.
Naylor, Paz Buenaventura (1975). 'Topic, focus and emphasis in the Tagalog verbal clause', *Oceanic Linguistics* 14: 12–79.
Naylor, Paz Buenaventura (1986), 'On the pragmatics of focus', in Paul Geraghty, Lois Carrington, and S. A. Wurm (eds.) *FOCAL 1: Papers from the Fourth International Conference on Austronesian Linguistics*, 43–57.
Nolasco, Ricardo Ma. (2003). 'Ang pagkaergatibo at pagkatransitibo ng mga wikang Pilipino: Isang Pagsusuri sa Sistemang Bose', PhD thesis, University of the Philippines, Diliman.
Otanes, Fe T. (1966). 'A contrastive analysis of English and Tagalog verb complementation', PhD thesis, UCLA.
Pallesen, Kemp A. (1985). *Culture contact and linguistic convergence*. Manila: Linguistic Society of the Philippines.
Pittman, Richard (1966). 'Tagalog -um- and mag-, an interim report'. Canberra: Linguistic Circle of Canberra Publications, ANU.
Porter, Doris (1977). *A Tboli grammar*. Manila: Linguistic Society of the Philippines.
Post, Ursula, and Mary Jane Gardner (1992). *Binukid dictionary*. Manila: Linguistic Society of the Philippines and Summer Institute of Linguistics.
Quakenbush, J. Stephen, Gail R. Hendrickson, and Josenita L. Edep (2010). *A brief overview of Agutaynen grammar*. Manila: SIL Publishing. Available at: www.sil.org/asia/philippines/plb_download.html
Reid, Lawrence A. (1992). 'On the development of the aspect system in some Philippine languages.' *Oceanic Linguistics* 31(1):65-91.
Reid, Lawrence A. (2002). 'Determiners, nouns, or what? Problems in the analysis of some commonly occurring forms in Philippine languages', *Oceanic Linguistics* 41: 295-309.
Reid, Lawrence A. (2007). 'Another look at the marking of plural personal noun constructions in Austronesian languages', *Oceanic Linguistics* 46(1): 232-252.
Reid, Lawrence A. (2017). 'Re-evaluating the position of Iraya among Philippine languages', in Hsiu-chuan Liao (ed.), *Issues in Austronesian Historical Linguistics*. JSEALS Special Publication No. 1, 23–47.
Reid, Lawrence A. and Hsiu-chuan Liao (2004). 'A brief syntactic typology of Philippine languages', *Language and Linguistics* 5(2): 433-490.
Revel-Macdonald, Nicole (1979). *Le Palawan (Philippines)*. Langues et Civilisations de l'Asie du Sud-Est et du Monde Insulindien Monograph #4. Paris: Centre National de la Recherche Scientifique.
Riesberg, Sonja (2014). *Symmetrical voice and Linking in Western Austronesian Languages*. Boston and Berlin: De Gruyter Mouton.
Ross, Malcolm (1995). 'Reconstructing Proto Austronesian verbal morphology: evidence from Taiwan' in Paul Jen-kuei Li, Dah-an Ho, Ying-kuei Huang, Cheng-hwa Tsang, and Chiu-yu Tseng (eds.), *Austronesian studies relating to Taiwan*. Taipei: Institute of History and Philology, Academia Sinica, 727–791.

Ross, Malcolm (2002). 'The history and transitivity of western Austronesian voice and voice marking', in Fay Wouk and Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics, 17-62.
Ross, Malcolm (2006). 'Reconstructing the case-marking and personal pronoun systems of Proto Austronesian', in Henry Yungli Chang, Lillian M. Huang, and Dah-an Ho (eds.), *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*. Taipei: Institute of Linguistics, Academia Sinica, 521–563.
Savage, Dale T. (1986). 'A Reconstruction of Proto-Southern Mindanaon', *Studies in Philippine Linguistics* 6: 181-223.
Scebold, Robert A. (2003). *Central Tagbanwa: a Philippine language on the brink of extinction: sociolinguistics, grammar, and lexicon*. Manila: Linguistic Society of the Philippines.
Schachter, Paul, and Fe T. Otanes (1972). *Tagalog Reference Grammar*. Berkeley, Los Angeles, and London: University of California Press.
Schachter, Paul (1976). 'The subject in Philippine languages: actor, topic, actor-topic, or none of the above. In Charles Li (ed.), *Subject and Topic*. New York: Academic Press, 491–518.
Starosta, Stanley, Andrew K. Pawley, and Lawrence A. Reid (1982). 'The evolution of focus in Austronesian' in Amran Halim, Lois Carrington, and S. A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics, vol. 2, Tracking the travellers*. Canberra: Pacific Linguistics, 145-170.
Stevens, Alan M. (1969). 'Case grammar in Philippine languages', Paper read at the Fifth Meeting of the Linguistic Society of America, San Francisco.
Van den Berg, Rene and David Mead (this volume). 'Sulawesi'.
Venago, Martin F. (1929). *Ang mga paring Pilipino sa kasaysayan ng Inang bayan*. Maynila: [s. n.].
Veronika Mattes (2014). *Types of reduplication: A case study of Bikol*. Berlin and Boston: De Gruyter Mouton.
Vivian M. Forsberg (1992). 'A Pedagogical Grammar of Tboli', *Studies in Philippine Linguistics* 9: 1-110.
Wang, Peter, Robert Hunt, Jeff McGriff, and Richard E. Elkins (2006). *The Grammar of Matigsalug Manobo*. Manila: SIL Philippines.
Wolfenden, Elmer (1961). *A Re-statement of Tagalog Grammar*. Manila: Summer Institute of Linguistics and the Institute of National Language.
Wolff, John U. (1966). *Beginning Cebuano I*. New Haven: Yale University Press.
Wolff, John U. (1973). 'Verbal inflection in Proto-Austronesian', in Andrew B. Gonzalez (ed.), *Parangal Kay Cecilio Lopez*. Manila: Linguistic Society of the Philippines, 71-91
Wu, Jiang (2017). 'Plural Words in Austronesian Languages: Typology and History', MA thesis, Leiden University.
Zobel, Eric (2005). 'Buol', in Alexander Adelaar and Nikolaus Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar*. London and New York: Routledge, 625-648.
Zorc, R. David (1971). 'Palawano notes'. SEALANG Zorc Papers. URL: http://sealang.net/archives/zorc/ (accessed 12/26/20).
Zorc, R. David (1974). 'Internal and External Relationships of the Mangyan Languages', *Oceanic Linguistics* 13: 561–600.
Zorc, R. David (1977). *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Canberra: Pacific Linguistics.

Zorc, R. David (2019). 'Klata/Giangan: A new Philippine subgroup', *The Archive: Special Publication* 16: 33-51.
Zorc, R. David, Jason Lobel, and William Hall (this volume). 'The historical linguistics of the Philippines.'
Zuraw, Kie (2000). 'Patterned Exceptions in Phonology', PhD thesis, UCLA.

APPENDIX 1. ABBREVIATIONS

ABS – absolutive
ADJ – adjective
ASP – aspect marker
AV – actor voice
BEG – begun aspect
CAUS – causative
CO – "co-"/sharing relation
CV – conveyance voice
DEF – definite
DEP – dependent mood
DET – determiner
DIST – distributive
DUR – durative
ERG – ergative
EXT – existential
FUT – future
GEN – genitive case
GER – gerund
HYP – hypothetical
IMMD – immediate future
INCH – inchoative
INDEF – indefinite
INTNS – intensive
IPFV – imperfective
ITER – iterative
LNK – linker
LOC – locative
LV – locative voice
MODER – moderate degree
NEG – negative
NM – noun marker
NMLZ – nominalizer
NOM – nominative case
OBL – oblique case
PL – plural marker
POT – potentive
PRETEND – pretendative
PFV – perfective

PROG – progressive
PV – patient voice
RCT – recent perfective
REAS – reason voice
RECP – reciprocal
RL – realis
SBJV - subjunctive
SOC – sociative
SPEC – specific
STA – stative
TOP – topic marker
TR – transitivity related
UV – undergoer voice
VRB – verbalizer

APPENDIX 2. LANGUAGES CITED

Agutaynen [agn]
Aklanon [akl]
Binukid/Western Bukidnon Manobo [bkd]
Buhi Bikol [bhk]
Central Subanen [syb]
Central Tagbanwa [tgt]
Cotabato Manobo [mta]
Hanunoo [hnn]
Hiligaynon [hil]
Iranun [ill]
Iraya [iry]
Legazpi Bikol [bcl]
Maguindanao [mdh]
Maranao [mrw]
Matigsalug Manobo [mbt]
Naga Bikol [bcl]
Northern Subanen [stb]
Palawano (Brooke's Point) [plw]
Samareño/Waray/Samar-Leyte [war]
Sarangani Manobo [mbs]
Tagalog (Manila and Batangas) [tgl]
Tagkaulo [klg]
Tboli [tbl]
Tiruray [tiy]
Toratan (Toratán) [rth]